

LOCUS CK230595 1157 bp mRNA linear EST 09-DEC-2003
DEFINITION ILLUMIGEN MCQ 942 Katze MNP2 Macaca mulatta cDNA 5' similar to human SERPINE2 (Hs.21858), mRNA sequence.
ACCESSION CK230595
VERSION CK230595.1 GI:39636768
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.
REFERENCE 1. (bases 1 to 1157)
Katze, M.G., Thomas, M., Korth, M., Iadonato, S.P. and Magness, C.L. Large-scale Rhesus Macaque cDNA Sequencing Unpublished (2003)
Contact: C. Magness
Illumigen Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagness@illumigen.com
Sequenced on 2003.10.16. 472 Q20 bases. Assemblies in contig w/ 66 member(s). Contig contains 63 (2.7%) lib members.
PCR PRIMERs
FORWARD: CCTCATTAAGGGAACAAA
BACKWARD: CACTATAGGCGAATGGGTA
Insert Length: 1157 Std Error: 0.00
Plate: CL000007 row: E column: 08
Seq primer: CCTCATTAAGGGAACAAA
POLYA=Yes.
Location/Qualifiers
1..1157
/organism="Macaca mulatta"
/mol_type="mRNA"
/strain="Indian"
/db_xref="taxon:9544"
/sex="male"
/dev_stage="newborn infant"
/lab_host="E. coli SOLR"
/clone_lib="Katze MNP2"
/note="Organ: placenta; Vector: Uni-ZAP XR; Site 1: EcoR
1: Site 2: Xho I; Created from Stratagene ZAP-cDNA
Synthesis kit (catalog #200400) and ZAP-cDNA Gigapack III
Gold Cloning Kit (Catalog #200450)"
ORIGIN
Query Match 76.0%; Score 387; DB 14; Length 1157;
Best Local Similarity 94.1%; Pred. No. 7.7e-84;
Matches 413; Conservative 0; Mismatches 25; Indels 1; Gaps 1;
QY 71 TGAATATGATTTGGGAGGAGATCAACATATAGGCTAGCCGAGGAGAGTGTACCA 130
DB 1 TGAATATGATTTGGGAGGAGATCAACATATAGGCTAGCCGAGGAGAGTGTACCA 60
QY 131 CTTCTGAGCTGATTTGGGCTTATGTTGGAGCCGGTGTAGAGAGCCTCAGCAAGAGAA 190
DB 61 CTTCTGAGCTGATTTGGGCTTATGTTGGAGCCGGTGTAGAGAGCCTCAGCAAGAGAA 120
QY 191 CCACCAACTGAAGTGGGATCTTGCACTGTGAGAGAGAGAGATCAGGCTGCA 250
DB 121 CCACCTAAGTGAAGTGGGATCTTGCACTGTGAGAGAGAGAGATCAGGCTGCA 180
QY 251 GCTGAGACTCAAGTGGCTGAGTGGAGCTGATCTCCAGAGCTGCTCAGTCAAGACT 310
DB 181 GCTGAGACTCAAGTGGCTGAGTGGAGCTGATCTCCAGAGCTGCTCAGTCAAGACT 240
QY 311 GGGGGTGAATGTGAAATGGTCTGTGATGACAGGGGAAGATTCTGCCAAATTCAGAACAA 370
DB 241 GGGGATGAATGCGGATGGTCTGTGATGACAGGGGAAGATTCTGCCAAATTCAGAACAA 300
QY 371 TTTAAATGCCAAGAGGTGACGGCAACACAGGTTTAAATGAAGACAGCTGAAC 430
DB 301 TTTAAATGCCAAGAGGTGACAGCAACACAGGTTTAAATGAAGACAGCTGAAC 360

QY 431 AACACAAACTGTTTATCTATCTAGATATTGACTTAAAAATATCAAAATACCTTTTCA 490
DB 361 AACAC-AAACTGTTTATCTATCTAAGATATTGACTTAAAAATATGAAATAAATCTTTGGC 419
QY 491 GCTTTCTCCAAAAA 509
DB 420 GCTTTCTCCAAAAA 438
D59091 394 bp mRNA linear EST 30-AUG-1995
HUM521C10B Clontech human placenta polyA+ mRNA (#6518) Homo sapiens
CDNA clone GEN-521C10 5', mRNA sequence.
D59091
ACCESSION D59091.1 GI:968725
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE EST.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1. (bases 1 to 394)
Fujiwara, T., Hirano, H., Katagiri, T., Kawai, A., Kuga, Y., Nagata, M., Okuno, S., Ozaki, K., Shimizu, F., Shimada, Y., Shinomiya, H., Takaichi, A., Takeda, S., Watanabe, T., Takahashi, E., Hirai, Y., Maekawa, H., Shin, S. and Nakamura, Y.
Fujiwara et al. (1995)
Unpublished (1995)
Contact: Tsutomu Fujiwara
Otsuka GEN Research Institute
463-10 Kagasuno Kawauchi-cho, Tokushima, 771-01 Japan
Tel: 0886-65-2888
Fax: 0886-37-1035.
Location/Qualifiers
1..394
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GEN-521C10"
/clone_lib="Clontech human placenta polyA+ mRNA (#6518)"
ORIGIN
Query Match 74.7%; Score 380.4; DB 14; Length 394;
Best Local Similarity 99.5%; Pred. No. 2.7e-82;
Matches 392; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 20 GTCGGCTTCTGCTGTCTGGACTTTTCTGTCCACTGAGAGCAGCTGTGTGAATATG 79
DB 1 GTCGGCTTCTGCTGTCTGGACTTTTCTGTCCACTGAGAGCAGCTGTGTGAATATG 60
QY 80 ATTGGGAGGAGATCAACATATAGGCTTAGGCCGAGGAGAGTGTACCACTCTCTGAG 139
DB 61 ATTGGGAGGAGATCAACATATAGGCTTAGGCCGAGGAGAGTGTGTACCACTCTCTGAG 120
QY 140 CTGATTTGGGCTTATGTTGGAGCCCGGTGTAGTGGAGCCTCAGCAAGAGAACCACT 199
DB 121 CTGATTTGGGCTTATGTTGGAGCCCGGTGTAGTGGAGCCTCAGCAAGAGAACCACT 180
QY 200 GAAAGTCGGATCTGCACTGTGTAGGAGAGAGAGATCAGGCTGAGACT 259
DB 181 GAAAGTCGGATCTGCACTGTGTAGGAGAGAGAGATCAGGCTGAGACT 240
QY 260 CAAGTGTCTGCTGAGAGCTGATCTCCAGAGCTGTCTCAGTCAAGAGCTGGGGTGA 319
DB 241 CAAGTGTCTGCTGAGAGCTGATCTCCAGAGCTGTCTCAGTCAAGAGCTGGGGTGA 300
QY 320 TGTGGAATGCTGCTGTATGACAGGGGAAGATTCTGCCAAATTCAGAACAAATTTAAATG 379
DB 301 TGTGGAATGCTGCTGTATGACAGGGGAAGATTCTGCCAAATTCAGAACAAATTTAAATG 360
QY 380 CCAGAA-GGAGGTGACAGGCAACCAAGGTTTAA 412

Db 361 CCAGAGGGAGGTGACAGGCAACACAGGTTAA 394
|||||
BX283272 382 bp mRNA linear EST 04-MAR-2003
LOCUS BX283272 NIH MGC 79 Homo sapiens cDNA clone IMAGE958D111343 ;
DEFINITION IMAGE:4605658, mRNA sequence.
ACCESSION BX283272
VERSION BX283272.1 GI:28847726
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 382)
Ebert,L., Heil,O., Hennig,S., Neubert,P., Partisch,B., Peters,M.,
Radelof,U., Schneider,D. and Korn,B.
Human Unigeneset - RZPD3
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGE958D111343.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-
bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
pCW-M13u, Primer sequence: CGTTGTAACGACGGCCACT.

FEATURES
source
Location/Qualifiers
1..382
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE958D111343 ; IMAGE:4605658"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NIH_MGC_79"
/note="Organ: Placenta; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggccgctggcc); Site_2: SfiI
(ggccattatggcc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3'
and 3' adaptor sequence:
5'-ATTCTAGCGCGCGCGGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.3
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 72.9%; Score 371; DB 13; Length 382;
Best Local Similarity 100.0%; Pred. No. 5.3e-80;
Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTGTGAGGGTGTGAGGTGCGGTTCTGCTGTCTGGACTTTTCTGTCTCCACTGAGA 60
Db 12 GAGTTGTGAGGGTGTGAGGGTTCGCTCTGCTGTCTGGACTTTTCTGTCTCCACTGAGA 71
Qy 61 CGCAGCTGTGTGAATATGATTTGGCGAGGAAGATCAACATATAGCCTAGGCCGAGGAG 120
Db 72 CGCAGCTGTGTGAATATGATTTGGCGAGGAAGATCAACATATAGCCTAGGCCGAGGAG 131
Qy 121 AAGTGTACCACTCTCTGAGCTGATTTGGCCCTATCTCTGAGCCCGGTGATGAGGAGCTCA 180
Db 132 AAGTGTACCACTCTCTGAGCTGATTTGGCCCTATCTCTGAGCCCGGTGATGAGGAGCTCA 191

Search completed: April 26, 2004, 01:58:56
Job time : 928.538 secs

Qy 181 GCAAGAGGAACCAACCACTGAAAAGTGGGATCTCTGCACCTGCTCAGGAGAGAGAAGA 240
Db 192 GCAAGAGGAACCAACCACTGAAAAGTGGGATCTCTGCACCTGCTCAGGAGAGAGAAGA 251
Qy 241 TCAGGGTGCAGCTCAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCA 300
Db 252 TCAGGGTGCAGCTCAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCA 311
Qy 301 GTCGAAGACTGGGGGTGAATGTGGAATGCTCCTGATGACCAAGGGAAGATTCTGCCAA 360
Db 312 GTCGAAGACTGGGGGTGAATGTGGAATGCTCCTGATGACCAAGGGAAGATTCTGCCAA 371
Qy 361 ATCAGAAACAAT 371
Db 372 ATCAGAAACAAT 382

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	483.4	95.0	659	4	US-09-620-312D-634	Sequence 634, App
2	214.4	42.1	767	4	US-09-621-975-1745	Sequence 1745, Ap
3	214.4	42.1	781	4	US-09-702-705-808	Sequence 808, App
4	214.4	42.1	781	4	US-09-736-457-808	Sequence 808, App
5	214.4	42.1	781	4	US-09-614-1243-808	Sequence 808, App
6	214.4	42.1	781	4	US-09-671-325-808	Sequence 808, App
7	214.4	42.1	781	4	US-09-589-184-808	Sequence 808, App
8	214.4	42.0	399	4	US-09-702-705-69	Sequence 69, Appl
9	214.4	42.0	399	4	US-09-736-457-69	Sequence 69, Appl
10	214.4	42.0	399	4	US-09-614-1243-69	Sequence 69, Appl
11	214.4	42.0	399	4	US-09-671-325-69	Sequence 69, Appl
12	214.4	42.0	399	4	US-09-589-184-69	Sequence 69, Appl
13	210.6	41.4	528	4	US-09-163-748C-1	Sequence 1, Appli
14	209.4	41.1	538	1	US-08-370-648-14	Sequence 14, Appl
15	209.4	41.1	538	2	US-08-531-662B-14	Sequence 14, Appl
16	209.4	41.1	538	3	US-08-669-161A-14	Sequence 14, Appl
17	209.4	41.1	538	3	US-08-602-039-14	Sequence 14, Appl
18	202.4	39.8	539	1	US-08-370-648-18	Sequence 18, Appl
19	202.4	39.8	539	2	US-08-531-662B-18	Sequence 18, Appl
20	202.4	39.8	539	3	US-08-669-161A-18	Sequence 18, Appl
21	202.4	39.8	539	3	US-08-602-039-18	Sequence 18, Appl
22	202.4	39.8	540	1	US-08-370-648-16	Sequence 16, Appl
23	202.4	39.8	540	2	US-08-531-662B-16	Sequence 16, Appl
24	202.4	39.8	540	3	US-08-669-161A-16	Sequence 16, Appl
25	202.4	39.8	540	3	US-08-602-039-16	Sequence 16, Appl
26	200.8	39.4	532	1	US-08-370-648-17	Sequence 17, Appl
27	200.8	39.4	532	2	US-08-531-662B-17	Sequence 17, Appl

QY 61 CGCAGCT-----GTGTGAAATATGATTGGCGAGAGATCAACATATAGG 106
Db 195 CGCAGCTGATTCTGTTTTCAGTGTGAAATATGATTGGCGAGAGATCAACATATAGG 254
QY 107 CTTAGGCGGAGGAGAGTGTACCACTCTGTAGCTGATTGGGCTATGTGAGGCCGGT 166
Db 255 CTTAGGCGGAGGAGAGTGTACCACTCTGTAGCTGATTGGGCTATGTGAGGCCGGT 314
QY 167 GATGAGGAGCTCAGCAAGAGGAGCAACCACTGAAAGTCGGGATCCTGCACCTGTGTACG 226
Db 315 GATGAGGAGCTCAGCAAGAGGAGCAACCACTGAAAGTCGGGATCCTGCACCTGTGTACG 374
QY 227 GAGAGAGAGAGATCGGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAGCTGATCTC 286
Db 375 GAGAGAGAGAGATCGGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAGCTGATCTC 434
QY 287 CAGGAGCTGTCTCAGTCAAAAGACTGGGGGTGAATGTGAAATGGTCTGATGACCAAGGGG 346
Db 435 CAGGAGCTGTCTCAGTCAAAAGACTGGGGGTGAATGTGAAATGGTCTGATGACCAAGGGG 494
QY 347 AGATTTCGCAAAATCAGAACCAATTTAAATGCCAGAGAGGTGACAGGCAACACAG 406
Db 495 AGATTTCGCAAAATCAGAACCAATTTAAATGCCAGAGAGGTGACAGGCAACACAG 554
QY 407 GTTAAATGAAGCAAGCTCAAAACACACAAACTGTTTTATCTAAGATATTGACTTA 466
Db 555 GTTAAATGAAGCAAGCTGAAACACACAAACTGTTTTATCTAAGATATTGACTTA 614
QY 467 AAATATCAAAATAAACTTTTGAGCTTCTCCAAAAA 509
Db 615 AAATATCAAAATAAACTTTTGAGCTTCTCCAAAAA 657

RESULT 2

US-09-621-976-1745
; Sequence 1745, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1745
; LENGTH: 767
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: 83..658
; NAME/KEY: sig_peptide
; LOCATION: 83..157
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.6999980926514
; OTHER INFORMATION: seq LCALLSNFCPSTT/VK
US-09-621-976-1745

Query Match 42.1%; Score 214.4; DB 4; Length 767;
Best Local Similarity 80.6%; Pred. No. 1.7e-52;
Matches 287; Conservative 0; Mismatches 66; Indels 3; Gaps 3;
QY 155 CTGAGGCCGCTGATGAGGAGCTTCAGCAAGAGAGCAACCAACTGAAAGTCGGGATCCT 214
Db 403 CAGGAGCCCAAGTAATGGAGAGCCCCCAAAAGAGAGCAACCAAGTCGGGATCCT 462
QY 215 GCACCTGTCAGGAGAGAGAGATCAGGGTCAGCTCAGCTCAGCTCAGCTCAGCTG 274
Db 463 ACACCTGGGCGAGAGAGAGATCAGGATACAGCTCAGCTCAGCTCAGCTCAGCTG 522

QY 275 GAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGAGCTGGGGGTCAATGTGGAAT -GGTCC 333
Db 523 GAAGCTGATCTGCAAGAGCTGCATCAGTCAACACCGGGATAAATCTGATTGGGTTTC 582
QY 334 TGATGACCAAGGAGAGATTTGCAAAATCAGAACCAATTTTAAATGCGCAAGAGGTGA 393
Db 583 CGGGCTCAAGGTGAAGATAATACCTTAAAGAGGAACTGTAAATGCCAAGACAGGTGA 642
QY 394 CAGCAACCAACAGGCTTTAAATGAAGCAAGCTGAAACCAACCAAACTGTTTTATCTAA 453
Db 643 AGACCAACCAACAGCTTTAAATGAAGCAAGCTGAAACCAACGC -AAGCTGGTTTTATATTA 701
QY 454 GATATTTGACTTTAAATATCAAAATCAAACTTTTGAGCTTCTCCAAAAA 509
Db 702 GATATTTGACTT-AACTATCTCAATAAGTTTTTGAGCTTTCACCAAAAAA 756

RESULT 3

US-09-702-705-808
; Sequence 808, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedrick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 808
; LENGTH: 781
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-808

Query Match 42.1%; Score 214.4; DB 4; Length 781;
Best Local Similarity 80.6%; Pred. No. 1.7e-52;
Matches 287; Conservative 0; Mismatches 66; Indels 3; Gaps 3;
QY 155 CTGGAGCCCGTGTATGAGGAGCTTCAGCAAGAGAGCAACCAACTGAAAGTCGGGATCCT 214
Db 426 CAGGAGCCCAAGTAATGGAGAGCCCCCAAAAGAGAGCAACCAAGTCGGGATCCT 485
QY 215 GCACCTGTCAGGAGAGAGAGATCAGGGTCAGCTCAGCTCAGCTCAGCTCAGCTG 274
Db 486 ACACCTGGGCGAGCAGACAGAGAGAGATCAGGATCAGCTCAGCTCAGCTCAGCTG 545
QY 275 GAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGAGACTGGGGGTGAATGTGGAAT -GGTCC 333
Db 546 GAAGTGTATCTGCAAGAGCTGCATCAGTCAACACCGGGGATAAATCTGATTGGGTTTC 605
QY 334 TGATGACCAAGGAGAGATTTCTGCCAAATCAAGAAATCAAGAAATTTTAAATGCCAAGAGGTGA 393
Db 606 CGGCTCAAGGTGAAGATAATACCTTAAAGAGAGCACTGTAAATGCCAAGAGAGGTGA 665
QY 394 CAGGCAACCAAGCTTTAAATGAAGAGAGAGCTGAAACCAACCAAACTGTTTTATCTAA 453
Db 666 AGAGCAACCAAGCTTTAAATGAAGAGAGAGCTGAAACCAACGC -AAGCTGGTTTTATATTA 724
QY 454 GATATTTGACTTTAAATATCAAAATCAAACTTTTGAGCTTCTCCAAAAA 509
Db 725 GATATTTGACTT-AACTATCTCAATAAGTTTTTGAGCTTTCACCAAAAAA 779

```
RESULT 4
US-09-736-457-808
; Sequence 808, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 808
; LENGTH: 781
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-808

Query Match 42.1%; Score 214.4; DB 4; Length 781;
Best Local Similarity 80.6%; Pred. No. 1.7e-52;
Matches 287; Conservative 0; Mismatches 66; Indels 3; Gaps 3;

QY 155 CTGGAGCCCGGTGATGAGGAGCCTTCAGTCAAGAGGAAACCAACCACTGAAAGTCGGATCCT 214
DB 426 CAGGAGCCCAAGTAATGGAGAGCCCCCAAAAGAGAACCCAGCAGCTGAAAGTCGGATCCT 485

QY 215 GCACCTGGTCAGGAGAGAGAGATCAGGGTCGAGCTGAGACTCAAGTCGCTGACCTG 274
DB 486 ACACCTGGGCGAGCAGACAGAGAGATCAGGATACAGCTGAGATCCCGAGTCGGACATG 545

QY 275 GAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGAGACTGGGGGTGAATGTGGAAT-GGTCC 333
DB 546 GAAGGTGATCTCCAGAGCTGCAATCAGTCAACACCCGGGGATAATCTGGATTGGGTTT 605

QY 334 TGATCACCAGGGGAAAGATTCTGCCAAAATCAGACAAATTTAAATGCCAGAGAGGTGA 393
DB 606 CGGCGTCAAGGTGAAGATAATACCTTAAAGAGAGACACTGTAAATGCCAGAGCGGTGA 665

QY 394 CAGGCAACCCACAGGTTTAAATGAACAGAGCTGAAACAACTGTTTATCTAA 453
DB 666 AGAGCAACCCACAGTTTAAATGAACAGAGCTGAAACAACTGTTTATCTAA 724

QY 454 GATATTGACTTTAAATAATATCAAAATTTTGCAGCTTTTCCAAAAA 509
DB 725 GATATTGACTT-AAACTATCTCAATAAAGTTTTCAGCTTTTCCAAAAA 779

RESULT 5
US-09-614-124B-808
; Sequence 808, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9

Query Match 42.1%; Score 214.4; DB 4; Length 781;
Best Local Similarity 80.6%; Pred. No. 1.7e-52;
Matches 287; Conservative 0; Mismatches 66; Indels 3; Gaps 3;

QY 155 CTGGAGCCCGGTGATGAGGAGCCTTCAGTCAAGAGGAAACCAACCACTGAAAGTCGGATCCT 214
DB 426 CAGGAGCCCAAGTAATGGAGAGCCCCCAAAAGAGAACCCAGCAGCTGAAAGTCGGATCCT 485

RESULT 6
US-09-671-325-808
; Sequence 808, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 808
; LENGTH: 781
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-671-325-808

Query Match 42.1%; Score 214.4; DB 4; Length 781;
Best Local Similarity 80.6%; Pred. No. 1.7e-52;
Matches 287; Conservative 0; Mismatches 66; Indels 3; Gaps 3;

QY 155 CTGGAGCCCGGTGATGAGGAGCCTTCAGTCAAGAGGAAACCAACCACTGAAAGTCGGATCCT 214
DB 426 CAGGAGCCCAAGTAATGGAGAGCCCCCAAAAGAGAACCCAGCAGCTGAAAGTCGGATCCT 485
```

215 GCACCTGGTCCAGGAGAGAGAGATCAGGGTGAGCTGAGACTCAAGTGCCTGACCTG 274
486 ACACCTGGGACAGACAGAGAGATCAGGATCAGCTGAGATCCAGTGGCGACATG 545
275 GAAGCTGATCTCAGGAGCTGTCTAGTCAAGACTGGGGGTGAATGTGGAAT-GGTCC 333
546 GAAGGTGATCTGCAAGAGCTGATCAGTCAACACCGGGGATAAATCTGGATTTGGGTTC 605
334 TGATGACCAAGGGAAGATTCTGCCAAATCAGAACTAAATTTAAATGCCAAGAGGAGTGA 393
606 CGCGTCAAGGTGAAGATTAATACCTAAAGAGGACACTGTAAATGCCAAGAGAGTGA 665
394 CAGGCAACACAGGTTTAAATGAGACAGCTGAAACACACAAAACCTGTTTTATCTAA 453
666 AGAGCAACACAGGTTTAAATGAGACAGCTGAAACACACG-AGCTGGTTTTATATTA 724
454 GATATTGACTTAAAAATATCAAAATATAAATTTTGCAGCTTTCTCCAAAAA 509
725 GATATTGACTT-AACTATCTCAATAAAGTTTTCAGCTTTTCCAAAAA 779

RESULT 7
US-09-589-184-808
; Sequence 808, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C8
; CURRENT APPLICATION NUMBER: US/09/589,184
; CURRENT FILING DATE: 2000-06-05
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 808
; LENGTH: 781
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-589-184-808

Query Match 42.1%; Score 214.4; DB 4; Length 781;
Best Local Similarity 80.8%; Pred. No. 1.7e-52;
Matches 287; Conservative 0; Mismatches 66; Indels 3; Gaps 3;

215 CTGAGCCCGGTGATGAGAGCTCTCAGCAAGAGGAAACCACTGAAAGTCCGGATCCT 214
426 CAGGAGCCAGTAATGGAGAGCCCAAAAGAGAACCAAGCTGAAAGTCCGGATCCT 485
215 GCACCTGGTCCAGGAGAGAGAGATCAGGGTGAGCTGAGACTCAAGTGCCTGACCTG 274
486 ACACCTGGGACAGACAGAGAGATCAGGATCAGCTGAGATCCAGTGGCGACATG 545
275 GAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGACTGGGGGTGAATGTGGAAT-GGTCC 333
546 GAAGGTGATCTGCAAGAGCTGATCAGTCAACACCGGGGATAAATCTGGATTTGGGTTC 605
334 TGATGACCAAGGGAAGATTCTGCCAAATCAGAACTAAATTTAAATGCCAAGAGGAGTGA 393
606 CGCGTCAAGGTGAAGATTAATACCTAAAGAGGACACTGTAAATGCCAAGAGAGTGA 665
394 CAGGCAACACAGGTTTAAATGAGACAGCTGAAACACACAAAACCTGTTTTATCTAA 453
666 AGAGCAACACAGGTTTAAATGAGACAGCTGAAACACACG-AGCTGGTTTTATATTA 724
454 GATATTGACTTAAAAATATCAAAATATAAATTTTGCAGCTTTCTCCAAAAA 509

725 GATATTGACTT-AACTATCTCAATAAAGTTTTCAGCTTTTCCAAAAA 779

RESULT 8
US-09-702-705-69
; Sequence 69, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 69
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-69

Query Match 42.0%; Score 214; DB 4; Length 399;
Best Local Similarity 80.3%; Pred. No. 1.6e-52;
Matches 286; Conservative 1; Mismatches 66; Indels 3; Gaps 3;

155 CTGAGCCCGGTGATGAGAGCTCTCAGCAAGAGGAAACCACTGAAAGTCCGGATCCT 214
44 CAGGAGCCAGTAATGGAGAGCCCAAAAGAGAACCAAGCTGAAAGTCCGGATCCT 103
215 GCACCTGGTCCAGGAGAGAGAGATCAGGGTGAGCTGAGACTCAAGTGCCTGACCTG 274
104 ACACCTGGGACAGACAGAGAGAGATCAGGATCAGCTGAGATCCAGTGGCGACATG 163
275 GAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGACTGGGGGTGAATGTGGAAT-GGTCC 333
164 GAAGGTGATCTGCAAGAGCTGATCAGTCAACACCGGGGATAAATCTGGATTTGGGTTC 223
334 TGATGACCAAGGGAAGATTCTGCCAAATCAGAACTAAATTTAAATGCCAAGAGGAGTGA 393
224 CGCGTCAAGGTGAAGATTAATACCTAAAGAGGACACTGTAAATGCCAAGAGGAGTGA 283
394 CAGGCAACACAGGTTTAAATGAGACAGCTGAAACACACAAAACCTGTTTTATCTAA 453
284 AGAGCAACCAAGTTTAAATGAGACAGCTGAAACACACG-AGCTGGTTTTATATTA 342
454 GATATTGACTTAAAAATATCAAAATATAAATTTTGCAGCTTTCTCCAAAAA 509
343 GATATTGACTT-AACTATCTCAATAAAGTTTTCAGCTTTTCCAAAAA 397

RESULT 9
US-09-736-457-69
; Sequence 69, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane

APPLICANT: Fan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 69
LENGTH: 399
TYPE: DNA
ORGANISM: Homo sapien
US-09-736-457-69

Query Match 42.0%; Score 214; DB 4; Length 399;
Best Local Similarity 80.3%; Pred. No. 1.6e-52;
Matches 286; Conservative 1; Mismatches 66; Indels 3; Gaps 3;
QY 155 CTGGAGCCCGGTGATGAGGAGCCTCAGCAAGAGGAGACCAACAACTGAAAGTCGGGATCCT 214
DB 44 CAGGAGCCCAAGTAATGGAGAGGCCCAAAAGAGAGACCAAGAGCTGAAAGTCGGGATCCT 103
QY 215 GCACCTGTGTCAGGAGAGAGAGATCAGGGTGCAGCTGAGACTCAAGTCGCTGACCTG 274
DB 104 ACACCTGGGCAGCAGACAGAGAGATCAGGATACAGCTGAGATCCCACTGGCGGACATG 163
QY 275 GAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGAGCTGGGGGTGAATGTGGAAT -GGTCC 333
DB 164 GAAGGTGATCTGCAGAGCTGCATCAGTCAACACCGGGGATAAATCTGGATTGGGTTTC 223
QY 334 TGATGACCGAGGGAAGTCTTGCACAAATCAGAACAAATTTAAATCCAGAGAGAGTGA 393
DB 224 CGCGGTCAAGGTGAAGATAATACCTAAAGAGGAAACACTGTAAATCCAGAGAGAGTGA 283
QY 394 CAGGCAACCAACAGGTTTAAATGAAGCAAGCTGAAACAAACAAAGCTGTTTTTATCTAA 453
DB 284 AGAGCAACCAACAGTTTAAATGAAGCAAGCTGAAACAAAGC -AAGCTGTTTTTATTA 342
QY 454 GATATTTGACTTAAATAATATCAAAATATAAATTTTTCAGCTTTTCCAAAAA 509
DB 343 GATATTTGACTT-AAACTATCTCAATAAAGTTTTTTCAGCTTTTCCACCAAAAAA 397

RESULT 10
US-09-614-124B-69
Sequence 69, Application US/09614124B
Patent No. 6630574
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C9
CURRENT APPLICATION NUMBER: US/09/614,124B
CURRENT FILING DATE: 2001-07-11
NUMBER OF SEQ ID NOS: 1668
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 69
LENGTH: 399
TYPE: DNA
ORGANISM: Homo sapien
US-09-614-124B-69

Query Match 42.0%; Score 214; DB 4; Length 399;
Best Local Similarity 80.3%; Pred. No. 1.6e-52;
Matches 286; Conservative 1; Mismatches 66; Indels 3; Gaps 3;

QY 155 CTGGAGCCCGGTGATGAGGAGCCTCAGCAAGAGGAGACCAACAACTGAAAGTCGGGATCCT 214
DB 44 CAGGAGCCCAAGTAATGGAGAGGCCCAAAAGAGAGACCAAGAGCTGAAAGTCGGGATCCT 103
QY 215 GCACCTGTGTCAGGAGAGAGAGATCAGGGTGCAGCTGAGACTCAAGTCGCTGACCTG 274
DB 104 ACACCTGGGCAGCAGACAGAGAGATCAGGATACAGCTGAGATCCCACTGGCGGACATG 163
QY 275 GAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGAGCTGGGGGTGAATGTGGAAT -GGTCC 333
DB 164 GAAGGTGATCTGCAGAGCTGCATCAGTCAACACCGGGGATAAATCTGGATTGGGTTTC 223
QY 334 TGATGACCGAGGGAAGTCTTGCACAAATCAGAACAAATTTAAATCCAGAGAGAGTGA 393
DB 224 CGCGGTCAAGGTGAAGATAATACCTAAAGAGGAAACACTGTAAATCCAGAGAGAGTGA 283
QY 394 CAGGCAACCAACAGGTTTAAATGAAGCAAGCTGAAACAAACAAAGCTGTTTTTATCTAA 453
DB 284 AGAGCAACCAACAGTTTAAATGAAGCAAGCTGAAACAAAGC -AAGCTGTTTTTATTA 342
QY 454 GATATTTGACTTAAATAATATCAAAATATAAATTTTTCAGCTTTTCCAAAAA 509
DB 343 GATATTTGACTT-AAACTATCTCAATAAAGTTTTTTCAGCTTTTCCACCAAAAAA 397

RESULT 11
US-09-671-325-69
Sequence 69, Application US/09671325
Patent No. 6667154
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C12
CURRENT APPLICATION NUMBER: US/09/671,325
CURRENT FILING DATE: 2000-09-26
NUMBER OF SEQ ID NOS: 1825
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 69
LENGTH: 399
TYPE: DNA
ORGANISM: Homo sapien
US-09-671-325-69

Query Match 42.0%; Score 214; DB 4; Length 399;
Best Local Similarity 80.3%; Pred. No. 1.6e-52;
Matches 286; Conservative 1; Mismatches 66; Indels 3; Gaps 3;
QY 155 CTGGAGCCCGGTGATGAGGAGCCTCAGCAAGAGGAGACCAACAACTGAAAGTCGGGATCCT 214
DB 44 CAGGAGCCCAAGTAATGGAGAGGCCCAAAAGAGAGACCAAGAGCTGAAAGTCGGGATCCT 103
QY 215 GCACCTGTGTCAGGAGAGAGAGATCAGGGTGCAGCTGAGACTCAAGTCGCTGACCTG 274
DB 104 ACACCTGGGCAGCAGACAGAGAGATCAGGATACAGCTGAGATCCCACTGGCGGACATG 163
QY 275 GAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGAGCTGGGGGTGAATGTGGAAT -GGTCC 333
DB 164 GAAGGTGATCTGCAGAGCTGCATCAGTCAACACCGGGGATAAATCTGGATTGGGTTTC 223
QY 334 TGATGACCGAGGGAAGTCTTGCACAAATCAGAACAAATTTAAATCCAGAGAGAGTGA 393
DB 224 CGCGGTCAAGGTGAAGATAATACCTAAAGAGGAAACACTGTAAATCCAGAGAGAGTGA 283

QY 394 CAGCAGCAGAGTCTTAATGAGACAGCTGAAACACACAAACTCTTTTATCTAA 453
Db 284 AGAGCAACCAAGTCTTAATGAGACAGCTGAAACACAGC-AAAGCTGGTTTATATTA 342
QY 454 GATATTTGACTTAAATATCAAAATAAATTTTGCAGCTTTCTCCAAAAA 509
Db 343 GATATTTGACTT-AAACTATCTCAATAAAGTTTTCAGCTTTTCCACCAAAAAA 397

RESULT 12

US-09-589-184-69
; Sequence 69, Application US/09589184
; Patent No. 6686447
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C8
; CURRENT APPLICATION NUMBER: US/09/589,184
; CURRENT FILING DATE: 2000-06-05
; NUMBER OF SEQ ID NOS: 827
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 69
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-589-184-69

Query Match 42.0%; Score 214; DB 4; Length 399;
Best Local Similarity 80.3%; Pred. No. 1.6e-52;
Matches 286; Conservative 1; Mismatches 66; Indels 3; Gaps 3;
QY 155 CTGAGCCCGGTGATGAGGAGCTCAGCAAGAGGACCACTGAAAGTCGGGATCT 214
Db 44 CAGGAGCCCAATGATGAGAGGCCCCCAAAAGAGAACCCAGCAGCTGAAAGTCGGGATCT 103
QY 215 GCACCTGCTCAGGAGAGAGAGATCAGGGTGCAGCTGAGACTCAAGTGCCTGACCTG 274
Db 104 ACACCTGGCAGCAGACAGAGAGATCAGATACAGCTGAGATCCAGTGCACATG 163
QY 275 GAAGCTGATCTCAGGAGCTGTCTCAGTCAAGCTGGGGGTGAATGTGGAAT-GGTCC 333
Db 164 GAAGGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGGATAATCTGGATTGGGTTT 223
QY 334 TGATGACAGGGGAGGATCTGCCAAATCAGAACATTTAAATGCCAGAGGAGTGA 393
Db 224 CGGCTCAAGTGAGATATATCTAAAGAGAACACTGTAAATGCCAGAGCAGTGA 283
QY 394 CAGCAACCAAGCTTTAAATGAAGACAGCTGAAACACAACTGTTTTTATCTAA 453
Db 284 AGAGCAACCAAGCTTTAAATGAAGACAGCTGAAACACAGC-AAAGCTGGTTTATATTA 342
QY 454 GATATTTGACTTAAATATCAAAATAAATTTTGCAGCTTTCTCCAAAAA 509
Db 343 GATATTTGACTT-AAACTATCTCAATAAAGTTTTCAGCTTTTCCACCAAAAAA 397

RESULT 13

US-09-163-748C-1
; Sequence 1, Application US/09163748C
; Patent No. 6509172
; GENERAL INFORMATION:
; APPLICANT: DeBacker, Oliver
; APPLICANT: Van den Eynde, Benoit
; APPLICANT: Boon-Failleur, Thierry
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Encode Member Of The Gage

; TITLE OF INVENTION: Family, The Proteins Encoded, And Uses Thereof
; FILE REFERENCE: LUD 5558
; CURRENT APPLICATION NUMBER: US/09/163,748C
; CURRENT FILING DATE: 1998-09-30
; NUMBER OF SEQ ID NOS: 27
; SEQ ID NO 1
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-163-748C-1

Query Match 41.4%; Score 210.6; DB 4; Length 528;
Best Local Similarity 69.4%; Pred. No. 1.8e-51;
Matches 353; Conservative 0; Mismatches 134; Indels 22; Gaps 4;
QY 20 GTCGGTTCCTGCTGCTGAGACTTTTCTGTCACACAGAGCAGCTGTGTGAAATATG 79
Db 16 GTGTGGTTCCTGCTGCTGAGACTTTTCTGTCACAGATTCATCTGTGTGAAATATG 75
QY 80 ATTTGGGAGGAGATCAACATATAGCCCTAGGCCGAGGAGAAAGTGTACCACCTCTCTGAG 139
Db 76 AGTTGGGAGGAGATCGACCTATCGGCCTAGACCAAGAGCCTACGTAGAGCCTCTCTGAA 135
QY 140 CTGATTGGGCTATGC-----TGGAGCCGCTGATGAGGAG-----CCTCAG 181
Db 136 ATGATTGGGCTATGCGGCCCGGAGCAGTTCAAGTATGAAGTGGAAACAGCAACCTGAA 195
QY 182 CAAGAGGAACCAACCACTGAAAGTCGGGATCTGACCTGTCAGGAGAGAGAAAGAT 241
Db 196 GAAGGGAACCAACCACTGAAAGTCGGGATCTGACCTGTCAGGAGAGAGAGAT 252
QY 242 CAGGTCAGCTGAGACTCAAGTGCCTGACCTGAGCTGATCTCCAGGAGCTCTCTCAG 301
Db 253 GAGGAGGATCTGCAAGTCAAGGCCCAAGCCTGAAGCTATATCCAGGAGAGAGGTCAC 312
QY 302 TCAAGAGCTGGGGTGAATGTGGAATGCTCCTGATGACCAAGGGAAGATTTCTCCAAA 361
Db 313 CCACAGACTGGGTGTGAGTGTGAAGATGGTCTCTGATGGCAGGAGATGGACCCCAAT 372
QY 362 TCAGACATTTAAATGCCAGAGGAGGTGACAGGCAACCAAGGTTTAAATGAAGACA 421
Db 373 CCAGAGGAGGTGAAACCGCTGAAAGAGGTGAAAGCAATCACAGTGTAAAGAGACA 432
QY 422 AGCTGAAACACACAAAACTGT-TTTTATCTAAATATTTGACTTTAAAAATATCAAAATA 480
Db 433 GCTTGAATGATGACAGGCTGCTCTCTATGTTGGAATTTGTTTCAATTAATTTCTCCCAATA 492
QY 481 AACTTTTCAGCTTTCTCCAAAAA 509
Db 493 AAGCTTTACAGCCTTCTGCAAAAGAAAAA 521

RESULT 14

US-08-370-648-14
; Sequence 14, Application US/08370648
; Patent No. 5648226
; GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Failleur, Thierry;
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM: Diskette, 5.25 inch, 360 kb storage
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370,648
FILING DATE: 10-JANUARY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250,162
FILING DATE: 27-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,039
FILING DATE: 22-JULY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5648226man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5323.2
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 538 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-370-648-14

Query Match 41.1%; Score 209; DB 1; Length 538;
Best Local Similarity 69.2%; Pred. No. 5.3e-51;
Matches 352; Conservative 0; Mismatches 135; Indels 22; Gaps 4;
QY 20 GTCGGCTTCCTGCTGCTGAGCTTTTCTGCTCCACTGAGAGCGAGCTGTGGAATATG 79
DB 27 GTGTGCTTCCTGCTGCTGAGCTTTTCTGCTCCACTGAGAGCGAGCTGTGGAATATG 86
QY 80 ATTTGCGGAGGAGATCAACATATAGCCCTAGCCGAGGAGAGTGTACCACTCTCTGAG 139
DB 87 AGTTGCGGAGGAGATCAACATATAGCCCTAGCCGAGGAGAGTGTACCACTCTCTGAG 146
QY 140 CTGATTTGGGCTTATGC-----TGGAGCCGGGTGATGAGGAG-----CCTCAG 181
DB 147 ATGATTTGGGCTTATGCGGCCGAGGAGCTTCTGATGATGAGTGGAAACACACCTGAA 206
QY 182 CAAGAGGAAACCAACAACTGAAAGTCTGCACTCTGCTGAGGAGAGAGAT 241
DB 207 GAAGGGGAAACCAACAACTCAACGCTCAGGATCTCTGAGCTGCTCAGGAG---GGAGAGGAT 263
QY 242 CAGGCTGACCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCAG 301
DB 264 GAGGAGGAGATCTGCAAGTCAAGGGCCGAGGAGCTGAGTGAAGCTCATAGCCAGGAAACAGGCTCAC 323
QY 302 TCAAGAGCTGGGGGTGAATGTGAATGCTGCTGATGACCAAGGAGAGATTTCTGCCAAA 361
DB 324 CCACAGACTGGGTGTGAGTGTGAAGTGTCTGATGGGCGAGGATGGAACCCGCCAAAT 383
QY 362 TCAGAACATTTAAATATGCCAGAGGAGGTGACAGGCAACCAACAGGTTTAAATGAAGACA 421
DB 384 CCAGAGGAGGTGAACAGCTGACAGAGGTGAAAGCAATCAACAGTGTGTAAGAGAGACA 443
QY 422 AGCTGAAACACACAAACTGT-TTTTATCTAAGTATTTGACTTAAATATCAAAATA 480
DB 444 CATTGAAATGATGCAAGGCTGCTTCTATGTTGGAATTTGTTTCAATTAATTTCTCCCAATA 503
QY 481 AACTTTTTCAGCTTTCTTCCAAAAA 509
DB 504 AAGCTTACAGCTTCTGCAAGAAAAA 532

RESULT 15
US-08-531-662B-14
; Sequence 14, Application US/08531662B
; Patent No. 5858689

GENERAL INFORMATION:
APPLICANT: Van der Bruggen, Pierre;
APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
APPLICANT: Boon-Palleur, Thierry
TITLE OF INVENTION: Isolated, Truncated Nucleic
TITLE OF INVENTION: Acid Molecules Which Code For GAGE Tumor
TITLE OF INVENTION: Rejection Antigen, The Tumor Rejection Antigen,
TITLE OF INVENTION: and Uses Thereof
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/531,662B
FILING DATE: 21-September-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/370,648
FILING DATE: 10-January-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250,162
FILING DATE: 27-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,039
FILING DATE: 22-July-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5858689man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5323.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 538 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-531-662B-14

Query Match 41.1%; Score 209; DB 2; Length 538;
Best Local Similarity 69.2%; Pred. No. 5.3e-51;
Matches 352; Conservative 0; Mismatches 135; Indels 22; Gaps 4;
QY 20 GTCGGCTTCCTGCTGCTGAGCTTTTCTGCTCCACTGAGAGCGAGCTGTGGAATATG 79
DB 27 GTGTGCTTCCTGCTGCTGAGCTTTTCTGCTCCACTGAGATTCATCTGTGGAATATG 86
QY 80 ATTTGCGGAGGAGATCAACATATAGCCCTAGCCGAGGAGAGTGTACCACTCTCTGAG 139
DB 87 AGTTGCGGAGGAGATCAACATATAGCCCTAGCCGAGGAGAGTGTACCACTCTCTGAG 146
QY 140 CTGATTTGGGCTTATGC-----TGGAGCCGGGTGATGAGGAG-----CCTCAG 181
DB 147 ATGATTTGGGCTTATGCGGCCGAGGAGCTTCTGATGATGAGTGGAAACACACCTGAA 206
QY 182 CAAGAGGAAACCAACAACTGAAAGTCTGCACTCTGCTGAGGAGAGAGAT 241
DB 207 GAAGGGGAAACCAACAACTCAACGCTCAGGATCTCTGAGCTGCTCAGGAG---GGAGAGGAT 263
QY 242 CAGGCTGACCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCAG 301
DB 264 GAGGAGGAGATCTGCAAGTCAAGGGCCGAGGAGCTGAGTGAAGCTCATAGCCAGGAAACAGGCTCAC 323

Qy	302	TCAAAGACTGGGGTGAATGTGGAATGGTCTCTGATGACCCAGGGGAAGATTCTGCCAAAA	361
Db	324	CCACAGACTGGGTGTGAGTGTGAGATGGTCTCTGATGGGCGAGAGATGGACCCGCCAAT	383
Qy	362	TCAGAACAAATTTAAATGCCAGAAAGGAGGTGACAGGCAACCAAGGTTTAAATGAAGACA	421
Db	384	CCAGAGGAGGTGAAGAACGCCCTGAAGAGGTGAAGCAATCACAGTGTAAAGAGAGACA	443
Qy	422	AGCTGAAACAAACACAAACTGT-TTTTATCTAAGATATTTGACTTAAAAATATCAAAATA	480
Db	444	CGTTGAATGATGCAGGCTGCTCTATGTTGGAAATTTGTTCAATTAATTTCTCCANTA	503
Qy	481	AACTTTGCAGCTTTCTCCAAAAA	509
Db	504	AAGCTTTACAGCCTTCTGCAAGAAAAA	532

Search completed: April 26, 2004, 02:10:45
Job time : 26.2508 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2004, 23:47:42 ; Search time 135,244 Seconds
(without alignments)
15988.422 Million cell updates/sec

Title: US-10-051-835-11
Perfect score: 509
Sequence: 1 gagtggaggtgtgaggg.....agcttttcccaaaaaaaa 509

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 337863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_24Jan04.*
1: Geneseqn1988s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002s.*
7: Geneseqn2003as.*
8: Geneseqn2003bs.*
9: Geneseqn2003cs.*
10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	509	100.0	509	7 ABX77605	Abx77605 Different
2	509	100.0	509	8 ACD42232	ACd42232 Human GAG
3	509	100.0	509	9 ADC24846	Adc24846 Human CDN
4	504.2	99.1	520	7 ABT15737	Abt15737 Human can
5	483.4	95.0	659	4 AAI58744	Aai58744 Human pol
6	483.4	95.0	659	8 ADB48724	Adb48724 Novel hum
7	480.6	94.4	532	4 AAI60530	Aai60530 Human pol
8	421.4	82.8	503	5 AAS69484	Aas69484 DNA encod
9	393.8	77.4	661	5 AAS69486	Aas69486 DNA encod
10	337.4	66.3	842	6 ABA93876	Aba93876 Human G p
11	337.4	66.3	673	4 AAK51918	Aak51918 Human pol
12	337.4	66.3	756	7 ABT15738	Abt15738 Human can
13	322.8	63.4	611	4 AAD14983	Aad14983 Human NOV
14	292.6	57.5	580	4 AAF59637	Aaf59637 Human cel
15	290.2	57.0	669	8 ACD42228	AcD42228 Human GAG
16	237.4	46.6	334	7 ABT15743	Abt15743 Human can
17	233	45.8	665	4 AAK52902	Aak52902 Human pol
18	214.4	42.1	626	7 ACC51056	Acc51056 Human bla
19	214.4	42.1	626	7 ABX76178	Abx76178 Lung can
20	214.4	42.1	626	7 ACC49534	Acc49534 Tumour-as
21	214.4	42.1	626	9 ADB75626	Adb75626 Prostate
22	214.4	42.1	637	6 ABX50912	Abx50912 XAGE-1 ge
23	214.4	42.1	762	5 AAH64751	Aah64751 Human sec

24	214.4	42.1	781	5 AAF68861	Aaf68861 Human lun
25	214.4	42.1	781	6 ABK38772	Abk38772 cDNA enco
26	214.4	42.1	781	7 ACA11101	AcA11101 Human lun
27	214.4	42.1	781	7 ACA02287	AcA02287 Lung can
28	214	42.0	399	5 AAF68151	Aaf68151 Human lun
29	214	42.0	399	6 ABK38062	Abk38062 cDNA enco
30	214	42.0	399	7 ACA10391	AcA10391 Human lun
31	214	42.0	399	7 ABX99342	Abx99342 Lung can
32	214	42.0	399	10 ADE72125	Ade72125 Human lun
33	211.4	41.5	463	6 ABA92217	Aba92217 Melanoma
34	210.6	41.4	528	7 ADA15801	Ada15801 Human GAG
35	209	41.1	538	2 AAX90519	Aax90519 GAGE-2 tu
36	209	41.1	1345	4 AAS60826	Aas60826 Human can
37	207	40.7	530	7 ABZ20463	Abz20463 GAGE-2 fu
38	205	40.3	541	2 AAV18719	Aav18719 cDNA enco
39	202.8	39.8	539	2 AAV18721	Aav18721 cDNA enco
40	202.4	39.8	540	2 AAX90523	Aax90523 GAGE-6 tu
41	202.4	39.8	540	2 AAX90521	Aax90521 GAGE-4 tu
42	201.8	39.6	535	2 AAV18717	Aav18717 cDNA enco
43	201.6	39.6	750	5 AAS91235	Aas91235 DNA encod
44	201.2	39.5	532	2 AAV18720	Aav18720 cDNA enco
45	200.8	39.4	532	2 AAX90522	Aax90522 GAGE-5 tu

ALIGNMENTS

RESULT 1
ABX77605
ID ABX77605 standard; cDNA; 509 BP.
XX
AC ABX77605;
XX
DT 09-APR-2003 (first entry)
XX
DE Differentially expressed breast cancer associated cDNA #100.
XX
KW Breast cancer; differential gene expression; BC-cDNA;
KW breast cancer diagnosis; breast cancer monitoring;
KW breast cancer treatment; breast cancer staging; gene; ss.
XX
OS Unidentified.
XX
PN US2002156263-A1.
XX
PD 24-OCT-2002.
XX
PF 04-OCT-2001; 2801US-00974298.
XX
PR 05-OCT-2000; 2000US-0238331P.
XX
PA (CHEN/) CHEN H.
XX
PI Chen H;
XX
DR WPI; 2003-182653/18.
XX
PT New cDNAs, which are differentially expressed in (metastatic) breast
PT cancer useful for diagnosing or staging, breast cancer, or for monitoring
PT the treatment of breast cancer in an individual.
XX
PS Claim 1; SEQ ID NO 126; 30pp; English.
XX
CC The invention describes a combination of cDNAs (designated BC-cDNAs),
CC which are differentially expressed in breast cancer. The combination
CC includes 152 cDNA sequences, or their complements. The protein encoded by
CC any of these BC-cDNAs is useful for screening several molecules or
CC compounds to identify at least one ligand that specifically binds the
CC protein, producing or preparing polyclonal or monoclonal antibodies, or
CC purifying antibodies from a sample. The antibodies, which specifically
CC bind the protein differentially expressed in breast cancer is useful for
CC detecting the expression of a protein in a sample. The BC-cDNAs are also
CC useful for diagnosing, monitoring the treatment of, or staging, breast

Qy	241	TCAGGCTGCAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCA	300
Db	241	TCAGGCTGCAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCA	300
Qy	301	GTCAAAAGACTGGGGGTGAATGTGGAATGTCTCTGATGACCCAGGGGAAGATTCTGCCAAA	360
Db	301	GTCAAAAGACTGGGGGTGAATGTGGAATGTCTCTGATGACCCAGGGGAAGATTCTGCCAAA	360
Qy	361	ATCAGAACATTTAAATGCCAGAGAGGTGACGCGACACACACAGTTTAAATGAAGAC	420
Db	361	ATCAGAACATTTAAATGCCAGAGAGGTGACGCGACACACACAGTTTAAATGAAGAC	420
Qy	421	AAGCTGAAACAACACACAAAACCTGTTTTTATCTAAGATATTGACTTAAAAATATCAAAATA	480
Db	421	AAGCTGAAACAACACACAAAACCTGTTTTTATCTAAGATATTGACTTAAAAATATCAAAATA	480
Qy	481	AACTTTTGCAGCTTTCTCCAAAAA	509
Db	481	AACTTTTGCAGCTTTCTCCAAAAA	509
RESULT 4			
ABT15737			
ID	ABT15737	standard; DNA; 520 BP.	
XX	AC	ABT15737;	
XX	28-MAR-2003	(first entry)	
DT	Human cancer/testis	antigen coding sequence - SEQ ID No 38.	
DE	Human; gene; ds;	gene therapy; vaccine; cancer; cancer/testis antigen;	
KW	CT antigen.		
XX	Homo sapiens.		
OS	WO200278526-A2.		
FN	10-OCT-2002.		
XX	29-MAR-2002;	2002WO-US009508.	
PF	30-MAR-2001;	2001US-0280718P.	
XX	20-APR-2001;	2001US-0285154P.	
FR	05-OCT-2001;	2001US-0327432P.	
XX	22-JAN-2002;	2002US-00054683.	
XX	(LUDW-) LUDWIG INST	CANCER RES.	
PA	(CORR) CORNELL RES	FOUND INC.	
PA	Old LJ, Scanlan MJ, Chen Y;		
PI	WPI; 2003-040608/03.		
XX	P-PSDS; ABJ19256.		
DR	Diagnosing cancer	comprises contacting a biological sample isolated from	
XX	a subject with an agent	that specifically binds to a nucleic acid	
PT	molecule, its expression	product or fragment or an antibody that binds to	
PT	the product or fragment.		
XX	Claim 7; Page 145; 155pp;	English.	
PS	The invention	comprises a method for diagnosing cancer, the method	
XX	involves detecting the	DNA or protein sequences of human cancer/testis	
CC	(CT) antigens that	are disclosed in the invention. The method of the	
CC	invention is useful	for detecting/diagnosing, treating and monitoring a	
CC	cancer or condition	characterised by the expression of a human CT	
CC	antigen. The present	DNA sequence encodes a human CT antigen of the	
XX	invention		
XX	Sequence 520 BP; 167 A; 100 C; 140 G; 113 T; 0 U; 0 Other;		


```
Query Match 99.1%; Score 504.2; DB 7; Length 520;
Best Local Similarity 99.4%; Pred. No. 2.4e-132;
Matches 506; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAGTTGTGAGGTGTGAGGTCGCGCTTCTGCTGTCTGAGCTTTTCTGTCCCACTGAGA 60
DB 1 GAGTTGTGAGGTGTGAGGTCGCGCTTCTGCTGTCTGAGCTTTTCTGTCCCACTGAGA 60
QY 61 CGCAGCTGTGTGAATATGATTTGGCGAGGAAGATCAACATATAGGCTAGGCCGAGGAG 120
DB 61 CGCAGCTGTGTGAATATGATTTGGCGAGGAAGATCAACATATAGGCTAGGCCGAGGAG 120
QY 121 AAGTGTACCACTCTGAGCTGATTTGGGCTATGCTGAGCGCGGTGATGAGGAGCTTCA 180
DB 121 AAGTGTACCACTCTGAGCTGATTTGGGCTATGCTGAGCGCGGTGATGAGGAGCTTCA 180
QY 181 GCAAGAGGAACACCAACTGAAAGTCGGATCCTGCACCTGCTCAGGAGAGAGAAGA 240
DB 181 GCAAGAGGAACACCAACTGAAAGTCGGATCCTGCACCTGCTCAGGAGAGAGAAGA 240
QY 241 TCAGGCTGAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCA 300
DB 241 TCAGGCTGAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCA 300
QY 301 GTCAAAGACTGGGGTGAATGTGGAATGGTCTGATGACCAAGGGAAGATTTGCCAAA 360
DB 301 GTCAAAGACTGGGGTGAATGTGGAATGGTCTGATGACCAAGGGAAGATTTGCCAAA 360
QY 361 ATCAGAACCAATTTAAATCCAGAGAGGAGTGCACAGGCAACACAGGTTTAAATGAAGAC 420
DB 361 ATCAGAACCAATTTAAATCCAGAGAGGAGTGCACAGGCAACACAGGTTTAAATGAAGAC 420
QY 421 AAGCTGAAACAAACACAAAACCTGTTTATCTAAGATATTTGACTTTAAAAATATCAAAATA 480
DB 421 AAGCTGAAACAAACCAAAACCTGTTTATCTAAGATATTTGACTTTAAAAATATCAAAATA 480
QY 481 AACTTTTGAGCTTTTCCAAAAA 509
DB 481 AACTTTTGAGCTTTTCCAAAAA 509

RESULT 5
AAI58744
ID AAI58744 standard; cDNA; 659 BP.
XX AC AAI58744;
XX DT 22-OCT-2001 (first entry)
XX DE Human polynucleotide SEQ ID NO 947.
XX KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX KW peripheral nervous system; neuropathy; central nervous system; CNS;
XX KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX KW leukaemia; ss.
XX OS Homo sapiens.
XX PN WC020153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WC-US034263.
XX PR 23-DEC-1999; 99US-00471275.
XX PR 21-JAN-2000; 2000US-00488725.
XX PR 25-APR-2000; 2000US-00552317.
XX PR 20-JUN-2000; 2000US-00598042.
XX PR 19-JUL-2000; 2000US-00620312.
XX PR 03-AUG-2000; 2000US-00653450.
XX PR 14-SEP-2000; 2000US-00662191.
```

```
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
XX PI Zhou P, Goodrich R, Drmanac RT;
XX XX WPI; 2001-442253/47.
XX DR P-PSDB; AAM39588.
XX PT Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.
XX PS Claim 1; SEQ ID NO 947; 10078pp; English.
XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
XX CC encoded polypeptides (AAM38642-AAM42213) with nontropic,
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide
XX CC of the invention may be used to treat diseases of the peripheral nervous
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and
XX CC localised neuropathies and central nervous system diseases, such as
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX CC utilisation of the activities such as: immune system suppression,
XX CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX CC assays for receptor activity, arthritis and inflammation, leukaemias and
XX CC C.N.S disorders. Note: The sequence data for this patent did not form
XX CC part of the printed specification
XX SQ Sequence 659 BP; 197 A; 137 C; 173 G; 152 T; 0 U; 0 Other;

Query Match 95.0%; Score 483.4; DB 4; Length 659;
Best Local Similarity 97.1%; Pred. No. 2.1e-126;
Matches 508; Conservative 0; Mismatches 1; Indels 14; Gaps 1;

QY 1 GAGTTGTGAGGTGTGAGGTCGCGCTTCTGCTGTCTGAGCTTTTCTGTCCCACTGAGA 60
DB 135 GAGTTGTGAGGTGTGAGGTCGCGCTTCTGCTGTCTGAGCTTTTCTGTCCCACTGAGA 194
QY 61 CGCAGCT-----GTGTGAATATATGATTTGGCGAGGAAGATCAACATATAGG 106
DB 195 CGCAGCTGTATTCTGTTTGCAGTGTGAATATGATTTGGCGAGGAAGATCAACATATAGG 254
QY 107 CTTAGGCCGAGGAGAGTGTACCACTCTGAGCTGATTTGGGCTATGCTGGAGCCGGT 156
DB 255 CTTAGGCCGAGGAGAGTGTACCACTCTGAGCTGATTTGGGCTATGCTGGAGCCGGT 314
QY 167 GATGAGGAGCTCAGCAAGAGGAACCAACCACTGAAAGTCGGATCCTGCACCTGGTCAG 226
DB 315 GATGAGGAGCTCAGCAAGAGGAACCAACCACTGAAAGTCGGATCCTGCACCTGGTCAG 374
QY 227 GAGAGAGAAGAGATCAGGCTGAGCTGAGCTCAAGTGCCTGACCTGGAGGCTGATTC 286
DB 375 GAGAGAGAAGAGATCAGGCTGAGCTGAGCTCAAGTGCCTGACCTGGAGGCTGATTC 434
QY 287 CAGGAGCTGTCTCAGTCAAAAGACTGGGGGTGAATGTGGAATGCTCTGATGACCAAGGG 346
DB 435 CAGGAGCTGTCTCAGTCAAAAGACTGGGGGTGAATGTGGAATGCTCTGATGACCAAGGG 494
QY 347 AAGATTCTGCCAAATCAGAACCAATTTAAATGCCAGAGAGGTGACAGGCAACCAAG 406
DB 495 AAGATTCTGCCAAATCAGAACCAATTTAAATGCCAGAGAGGTGACAGGCAACCAAG 554
QY 407 GTTTAAATGAGCAAGCTGAACCAACCAAACTGTTTTTATCTAAGATATTGACTTA 466
DB 555 GTTTAAATGAGCAAGCTGAACCAACCAAACTGTTTTTATCTAAGATATTGACTTA 614
QY 467 AAAATATCAAAATAAACTTTTGCAGCTTTTCCAAAAA 509
```


CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

XX SQ Sequence 532 BP; 154 A; 108 C; 150 G; 120 T; 0 U; 0 Other;
Query Match 94.4%; Score 480.6; DB 4; Length 532;
Best Local Similarity 98.0%; Pred. No. 1.2e-125;
Matches 497; Conservative 0; Mismatches 9; Indels 1; Gaps 1;
QY 1 GAGTGTGAGGGTGTGAGGTCGCGTCTCTGCTGTGAGCTTTTCTCTCCACTGAGA 60
DB 26 GAGTGTGAGGGTGTGAGGTCGCGTCTCTGCTGTGAGCTTTTCTCTCCACTGAGA 85
QY 61 CGCAGCTGTCTGAAATATGATTGGCGAGAGATCAACATATAGGCCCTAGGCCGAGGAG 120
DB 86 CGCAGCTGTCTGAAATATGATTGGCGAGAGATCAACATATAGGCCCTAGGCCGAGGAG 145
QY 121 AAGTGTACCACTCTGAGCTGATGGGCTATGCTGGAGCCCGGTGATGAGGAGCCTCA 180
DB 146 AAGTGTACCACTCTGAGCTGATGGGCTATGCTGGAGCCCGGTGATGAGGAGCCTCA 205
QY 181 GCAGAGGAGACCACTGAGAGTGGGATGCTGGAGCCCGGTGATGAGGAGCCTCA 240
DB 206 GCAGAGGAGACCACTGAGAGTGGGATGCTGGAGCCCGGTGATGAGGAGCCTCA 265
QY 241 TCAGGGTGACGTGAGACTCAAGTGCCTGACCTGGAGCTGATCTCCAGGAGCTGTCTCA 300
DB 266 TCAGGGTGACGTGAGACTCAAGTGCCTGACCTGGAGCTGATCTCCAGGAGCTGTCTCA 325
QY 301 GTCAAGAGCTGGGGTGAAATGTGGAAATGCTCTGATGACCGGGGAGAACTTCCCAA 360
DB 326 GTCAAGAGCTGGGGTGAAATGCGGAGATGCTCTGATGACCGGGGAGAACTTCCCAA 385
QY 361 ATCAGAACAAATTTAAATGCCAGAA-GGAGGTGACAGGCAACCCACAGGTTTAAATGAAGA 419
DB 386 GTCAAGAGCAATTTAAATGCCAGAGGAGTGCAGAGCAACCCACAGGTTTAAATGAAGA 445
QY 420 CAAGCTGAAACACACAAACTGTTTATCTAAGATATTTGACTTTAAATATCAAAAT 479
DB 446 CAAGCTGAAACACACAAACTGTTTATCTAAGATATTTGACTTTAAATATCAAAAT 505
QY 480 AAACCTTTTGCAGCTTTCTCCAAAAAAA 506
DB 506 AAACCTTTTGCAGCTTTCTCCAAAAAAA 532

RESULT 8

AA569484
ID AA569484 standard; cdna; 503 BP.

XX AC AA569484;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #5288.

XX DE Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX PN W0200175067-A2.

XX XX

PD 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US0008631.
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR P-ESDB; ABG05297.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 1; SEQ ID NO 5288; 103pp; English.
PS The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 503 BP; 158 A; 98 C; 121 G; 105 T; 0 U; 21 Other;

Query Match 82.8%; Score 421.4; DB 5; Length 503;

Best Local Similarity 98.6%; Pred. No. 6.4e-109;
Matches 425; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 69 TGTGAAATATGATTTGGCGAGGAAGATCAACATATATAGGCTAGCGGAGAGAGTGTAC 128

DB 52 TGTGAAATATGATTTGGCGAGGAAGATCAACATATATAGGCTAGCGGAGAGAGTGTAC 111

QY 129 CACCTCTGAGCTGATTTGGGCTATGCTGGAGCCGCTGATGAGGAGCTCAGCAAGAGG 188

DB 112 CACCTCTGAGCTGATTTGGGCTATGCTGGAGCCGCTGATGAGGAGCTCAGCAAGAGG 171

QY 189 AACCAACCACTGAAAGTCGGATCCTGCACCTGGTCAGGAGAGAGAAAGATCAGGGTG 248

DB 172 AACCAACCACTGAAAGTCGGATCCTGCACCTGGTCAGGAGAGAGAAAGATCAGGGTG 231

QY 249 CAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCAGTCAAAGA 308

DB 232 CAGTGAAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCAGTCAAAGA 291

QY 309 CTGGGGGTGAATGTGGAATGTGTCCTGATCAGCAGGGGAGAGATTTCCCAAAATCAGAAC 368

DB 292 CTGGGGGTGAATGTGGAATGTGTCCTGATCAGCAGGGGAGAGATTTCCCAAAATCAGAAC 351

QY 369 AATTTAAATGCCAGAGAGGAGGTGACAGGCAACACAGGTTTAAATGAGACAGAGCTGA 428

DB 352 AATTTAAATGCCAGAGAGGAGGTGACAGGCAACACAGGTTTAAATGAGACAGAGCTGA 411

QY 429 ACAACACAAAAGCTGTTTATCTTAAGATATTTGACTTAAATAATATCAAAATAAATCTTTTG 488
 |||||
 DB 412 ACAACCCCAAAAGCTGTTTATCTTAAGATATTTGACTTAAATAATATCGAAATAAATCTTTTG 471
 |||||
 QY 489 CAGCTTTCTCC 499
 |||||
 DB 472 CAGCTTTCTCC 482

RESULT 9

AA569486
 ID AA569486 standard; cDNA; 661 BP.

AC AA569486;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #5290.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

OS WO200175067-A2.

PN 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-839362/73.

DR P-PSDB; ABG05299.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 1; SEQ ID NO 5290; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AA564197-AA594564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 661 BP; 212 A; 129 C; 158 G; 162 T; 0 U; 0 Other;

Query Match 77.4%; Score 393.8; DB 5; Length 661;

Best Local Similarity 98.7%; Pred. No. 4.7e-101;

Matches 439; Conservative 0; Mismatches 2; Indels 4; Gaps 4;
 QY 65 GCTGTGTGAAATATGATTTGGCGAGGAAGATCAATATAGCCTAGGCCGAGGAGAACT 124
 |||||
 DB 217 GCAGTGTGAAATATGATTTGGCGAGGAAGATCAATATAGCCTAGGCCGAGGAGAACT 276
 |||||
 QY 125 GTACCACCTCCTGAGCTGATTTGGGCTATGCTGGAGCCCGGTGATGAGGAGCCTCAGCAA 184
 |||||
 DB 277 GTACCACCTCCTGAGCTGATTTGGGCTATGCTGGAGCCCGGTGATGAGGAGCCTCAGCAA 336
 |||||
 QY 185 GAGGAACCACTCACTGAAAGTGGGATCCTGACCT-GCTCAGGAGAGAGAGAGATCA 243
 |||||
 DB 337 GAGGAACCACTCACTGAAAGTGGGATCCTGACCTGCGGTGAGGAGAGAGAGATCA 396
 |||||
 QY 244 GGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAA-GCTGATCTCCAGGAGCTGTCTCAGT 302
 |||||
 DB 397 GGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAGGCTGATCTCCAGGAGCTGTCTCAGT 456
 |||||
 QY 303 CAAAGCTGGGGGTGAATGTGGAATGGTCTCTGATGACCCAGGGGAGATTTGCCAAAT 362
 |||||
 DB 457 CAAAGCTGGGGGTGAATGTGGAATGGTCTCTGATGACCCAGGGGAGATTTGCCAAAT 516
 |||||
 QY 363 CAGAACAA-TTTTAAATGCCAGAA-GGAGTGCAGGCAACACACAGGTTTAAATGAAGAC 420
 |||||
 DB 517 CAGAACAA-TTTTAAATGCCAGAGGAGGTGACAGGCAACACACAGGTTTAAATGAAGAC 576
 |||||
 QY 421 AAGCTGAAACCAACACAAAACCTGTTTTTATCTTAAGATATTGACTTAAAAATATCAAAATA 480
 |||||
 DB 577 AAGCTGAAACCAACACAAAACCTGTTTTTATCTTAAGATATTGACTTAAAAATATCGAAATA 636
 |||||
 QY 481 AACTTTTGCAGCTTCTCCAAAAA 505
 |||||
 DB 637 AACTTTTGCAGCTTCTCCAAAAA 661

RESULT 10

ABA93876

ID ABA93876 standard; cDNA; 642 BP.

AC ABA93876;

DT 07-MAY-2002 (first entry)

DE Human G protein-coupled receptor NOV4 encoding cDNA SEQ ID NO:9.

XX Human; G protein-coupled receptor; receptor; receptor; GPCR; NOVX; cardiant;
 KW antiarteriosclerotic; cytosolic; immunomodulatory; anorectic; antiviral;
 KW antiinflammatory; coagulant; gene therapy; cardiomyopathy; hypertension;
 KW atherosclerosis; valve disease; obesity; prostate cancer; adenocarcinoma;
 KW uterus cancer; fertility; haemophilia; immunodeficiency; AIDS;
 KW graft versus host disease; Crohn's disease; multiple sclerosis; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 207..542

XX FT /*tag= a

XX FT /product= "NOV4 protein"

XX WO200200691-A2.

XX 03-JAN-2002.

XX 27-JUN-2001; 2001WO-US020510.

XX 27-JUN-2000; 2000US-0214759P.

PR 31-OCT-2000; 2000US-0244546P.

PR 13-NOV-2000; 2000US-0248153P.

PR 11-JAN-2001; 2001US-0261014P.

XX 22-JAN-2001; 2001US-0263215P.

XX (CURA-) CURAGEN CORP.

KW Parkinson's disease; infertility; autoimmune disease; arthritis;
KW multiple sclerosis; allergy; wound healing; cytostatic; nootropic;
KW immunosuppressive; neuroprotective; vulnerary; hepatotropic; ds.
XX Homo sapiens.

XX Key Location/Qualifiers
FH 1. .173
5'UTR /tag= a
CDS 174..521
/tag= b
/product= "Human NOV4 protein"
3'UTR 522..611
/tag= c

WO200161009-A2.

23-AUG-2001.

15-FEB-2001; 2001WO-US004928.

15-FEB-2000; 2000US-0182723P.

15-FEB-2000; 2000US-0182724P.

15-FEB-2000; 2000US-0182733P.

22-FEB-2000; 2000US-0183896P.

23-FEB-2000; 2000US-0184275P.

23-FEB-2000; 2000US-0184482P.

23-FEB-2000; 2000US-0184497P.

24-FEB-2000; 2000US-0184744P.

13-APR-2000; 2000US-0197083P.

10-AUG-2000; 2000US-0223405P.

18-SEP-2000; 2000US-0233340P.

27-SEP-2000; 2000US-0236060P.

02-JAN-2001; 2001US-0259414P.

18-JAN-2001; 2001US-0262454P.

14-FEB-2001; 2001US-00783429.

(CURA-) CURAGEN CORP.

Malyankar UM, Tchernev VT, Padigar M, Taupier RJ, Spytek KA;

Majumder K, Guo X, Spaderna SK, Boldog FL;

WPI; 2001-514775/56.

P-PSDB; AAE08583.

Isolated novel polypeptides useful for diagnosis of and treating cancer, infertility, autoimmune diseases, arthritis, multiple sclerosis, allergies, wound healing and hepatic disorders.

Claim 9; Page 14; 140pp; English.

The present sequence is a human NOV4 DNA. The NOVX protein has homology with one of G-antigen (GAGE)-like protein, interferon, G-protein coupled receptor (GPCR), hepatocyte nuclear factor or mast cell protease. The NOVX is useful for treating or preventing a pathology associated with NOVX. It is also useful for determining the presence or amount of NOVX DNA in a sample, for identifying a potential therapeutic agent and in gene therapy. It is also useful for determining the presence of or predisposition to a disease associated with altered levels of NOVX. It is also useful for the diagnosis and treatment of proliferative disorders, e.g., cancer, immune disorders, hepatic disorders, e.g., cirrhosis, viral infections, e.g., hepatitis, neurotoxicity, system-related disorders, neurological disorders, e.g., Parkinson's disease, infertility, autoimmune diseases, arthritis, multiple sclerosis, allergies and wound healing

Sequence 611 BP; 196 A; 127 C; 152 G; 135 T; 0 U; 1 Other;

Query Match 63.4%; Score 322.8; DB 4; Length 611;

Best Local Similarity 79.5%; Pred No. 5, 5e-81;

Matches 435; Conservative 0; Mismatches 73; Indels 39; Gaps 3;

Db 62 GAGCTGTGAGAGTGTGAGGACACGTTCCAGCGTCTGAGACTTTCTCTCCCTACTGAGA 121
QY 61 CGCAGCTGT-----GTGAAATATGATTTG 84
Db 122 CGCAGCTATAGTTCGCGAGCCAGTCTCCAGGAACCTGAAATATGATTTG 181
QY 85 GCGAGGAGATCAACATATAGGCTAGGCGAGGAGAGTGTACCACTCTCTCAGCTGAT 144
Db 182 GCGAGGAGATCAACATATAGGCTAGGCGAGGAGAGTGTACCACTCTCTCAGCTGAT 241
QY 145 TGGGCTTATGCTGGAGCCCGGTGATGAGGAGCCTCAGCAAGAGAGAAACCACTGAAAG 204
Db 242 TGGGCTTATGCTGGAGCCCGGTGATGAGGAGCCTCAGCAAGAGAGAAACCACTGAAAG 301
QY 205 TCGGATCTTCGACCTG--GTCAGGAGAGAGAGATCAGGTCAGGTCAGCTGAGACTCAA 262
Db 302 TCGGATCTTCGACCTG--GTCAGGAGAGAGAGATCAGGTCAGGTCAGCTGAGACTCAA 361
QY 263 GTGCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGAGCTGGGGTGAATGT 322
Db 362 GTGCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGAGCTGGGGTGAATGT 421
QY 323 GGAATGCTCTGATGACCGAGGAGAGATTTCTCCAAATCAGACAAATTTAAATGCCA 382
Db 422 GAAGTGTGCTACTGATGTCAAGGGGAGATTTCTCCAAAGCAGAGCACTTTAAATGCCA 481
QY 383 GAAGGAGTGCAGGCAACCAACAGCTTTAAATGAAGCAAGCTGAAACACACAACTG 442
Db 482 GAAGCAGTGAAGGAAATCAGAGTTTAAAGGAAGATAGCTGAAACACACAACTG 540
QY 443 TTTTATCTAAGATATTGACTTAAATATCAAAATTAATCTTTGAGCTTTCTCCAAA 502
Db 541 TTTTATCTAAGATATTGACTTAAATATCAAAATTAATCTTTGAGCTTTCTCCAAA 600
QY 503 AAAAAA 509
Db 601 AAAAAA 607

RESULT 14

AAF59637

ID AAF59637 standard; cDNA; 580 BP.

XX AAF59637;

XX 24-APR-2001 (first entry)

XX Human cell cycle and proliferation protein CCVPR-48 cDNA, SEQ ID NO:102.

XX Cell cycle and proliferation protein; CCVPR; human; agonist; antagonist;
XX gene therapy; detection; gene therapy; transgenic animal disease model;
XX immune disorder; developmental disorder; cell signalling disorder;
XX cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
XX arteriosclerosis; asthma; allergy; diabetes mellitus;
XX menstrual cycle disorder; bacterial infection; ss.

OS Homo sapiens.

XX WO200107471-A2.

XX 01-FEB-2001.

XX 21-JUL-2000; 2000WO-US019948.

XX 21-JUL-1999; 99US-0145075P.

XX 08-SEP-1999; 99US-0153129P.

XX 10-NOV-1999; 99US-0164647P.

XX (INCY-) INCYTE GENOMICS INC.

XX Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;

XX Azimzai Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;

Qy	304	AAAGACTGGGGTCAATGTGAAATGGTCCTGATGACACAGGGGAAGATTCTGCCAAAATC	363
Db	472	AAAGACTGGGGATGGATGTGAAGTGGTACTGATGTCAAGGGGAAGATTCTACCAAAGC	531
Qy	364	AGAACAAATTAAATGCCAGAGGAGGTGACAGGCCACACAGGTTTAATGACACACRAG	423
Db	532	AGAGCACTTTAAATGCCAGAGCAGGTGAAGGGAATCACAGGTTTAAAGGAGATAAG	591
Qy	424	CTGAACAACACACAAAACCTGTTTTATCTTAAGATATTTGACTTAAATAATATCAAAATAAC	483
Db	592	CTGAACAACAC-AAACTGTTTTATATTAGATATTTTACTTTAAAGACTCTTAATAAT	650
Qy	484	TTTTG	488
Db	651	TTTTG	655

Search completed: April 25, 2004, 08:45:40
Job time : 139.244 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2004, 23:54:27 ; Search time 1340.3 Seconds
(without alignments)
16460.143 Million cell updates/sec

Title: US-10-051-835-11

Perfect score: 509

Sequence: 1 gagtgtgaggtgtgaggg.....agcttttcccaaaaaaaa 509

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_hgt_hum.*

31: em_hgt_inv.*

32: em_hgt_other.*

33: em_hgt_mus.*

34: em_hgt_pln.*

35: em_hgt_rod.*

36: em_hgt_mam.*

37: em_hgt_vrt.*

38: em_sy.*

39: em_hgt_hum.*

40: em_hgt_mus.*

41: em_hgt_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	505.8	99.4	548	9	BC062680	BC062680 Homo sapi
2	483.4	95.0	659	6	AR339143	AR339143 Sequence
3	457.8	89.9	550	6	AX921769	AX921769 Sequence
4	456	89.6	494	6	AX921771	AX921771 Sequence
5	439.4	86.3	493	9	HSX11881	AJ318881 Homo sapi
6	337.4	66.3	620	9	HSX11880	AJ318880 Homo sapi
7	337.4	66.3	639	9	BC009232	BC009232 Homo sapi
8	337.4	66.3	642	6	AX359705	AX359705 Sequence
9	322.8	63.4	611	6	AX226501	AX226501 Sequence
10	294.2	57.8	20587	9	HS118526A	268274 Human DNA s
11	294.2	57.8	62493	9	HS1193G15	AL117391 Human DNA
12	294.2	57.8	62493	9	HS1193G15	AL096838 Human DNA
13	294.2	57.8	145151	9	AL645949	AL645949 Human DNA
14	292.6	57.5	580	6	AX078298	AX078298 Sequence
15	214.4	42.1	529	9	BC009538	BC009538 Homo sapi
16	214.4	42.1	626	9	AF251237	AF251237 Homo sapi
17	214.4	42.1	637	6	AX455517	AX455517 Sequence
18	214.4	42.1	740	9	HSX118878	AJ318878 Homo sapi
19	214.4	42.1	767	6	AR414108	AR414108 Sequence
20	214.4	42.1	767	6	BD109661	BD109661 EST and e
21	214.4	42.1	781	6	AR273067	AR273067 Sequence
22	214.4	42.1	781	6	AR276648	AR276648 Sequence
23	214.4	42.1	781	6	AR406923	AR406923 Sequence
24	214.4	42.1	781	6	AX063181	AX063181 Sequence
25	214.4	42.1	781	6	AX368098	AX368098 Sequence
26	214	42.0	399	6	AR272357	AR272357 Sequence
27	214	42.0	399	6	AR275938	AR275938 Sequence
28	214	42.0	399	6	AR406213	AR406213 Sequence
29	214	42.0	399	6	AX062442	AX062442 Sequence
30	214	42.0	399	6	AX367359	AX367359 Sequence
31	212.4	41.7	463	9	HSX290447	AJ290447 Homo sapi
32	211.4	41.5	463	6	AX370580	AX370580 Sequence
33	210.6	41.4	528	6	AR275665	AR275665 Sequence
34	210.6	41.4	528	9	AF055473	AF055473 Homo sapi
35	210.6	41.4	551	9	BC018052	BC018052 Homo sapi
36	209	41.1	538	6	AR028488	AR028488 Sequence
37	209	41.1	538	6	BD231797	BD231797 Isolated
38	209	41.1	538	6	I55851	I55851 Sequence 14
39	209	41.1	538	6	BD132467	BD132467 Isolated,
40	209	41.1	1245	6	AX285022	AX285022 Sequence
41	207	40.7	530	9	HSU19143	U19143 Human GAGE-
42	202.4	39.8	539	6	AR028492	AR028492 Sequence
43	202.4	39.8	539	6	BD231801	BD231801 Isolated
44	202.4	39.8	539	6	I55855	I55855 Sequence 18
45	202.4	39.8	539	6	BD132471	BD132471 Isolated,

ALIGNMENTS

RESULT 1
BC062680 548 bp mRNA linear PRI 11-DEC-2003
LOCUS BC062680 Homo sapiens cDNA clone MGC:71925 IMAGE:4619720, complete cds.
DEFINITION BC062680
ACCESSION BC062680.1 GI:38541861
VERSION BC062680.1
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 548)
REFERENCE
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,

195 CGACGCTGATTTCTGTTGCGAGTGAATATGATTTGGCGAGGAGATCAACATATAGG 254
Db
107 CCTAGGCCGAGAGAGTGTACACCTCTGAGCTGATTTGGGCTATGCTGGAGCCGGT 166
Qy
255 CCTAGGCCGAGAGAGTGTACACCTCTGAGCTGATTTGGGCTATGCTGGAGCCGGT 314
Db
167 GATGAGGAGCTTCAGCAAGAGGAAACCAACAACTGAAAGTCCGGATCCTGCACTGGTGCAG 226
Qy
315 GATGAGGAGCTTCAGCAAGAGGAAACCAACAACTGAAAGTCCGGATCCTGCACTGGTGCAG 374
Db
227 GAGAGAGAGAGATGACGGGTGACGCTGAGACTCAAGTCCCTGACCTGGAGAGCTGATCTC 286
Qy
375 GAGAGAGAGAGATGACGGGTGACGCTGAGACTCAAGTCCCTGACCTGGAGAGCTGATCTC 434
Db
287 CAGGAGCTGTCTCAGTCAAGAGTGGGGTGAATGCGAAATGGTCTGTGATGACCCAGGG 346
Qy
435 CAGGAGCTGTCTCAGTCAAGAGTGGGGTGAATGCGAAATGGTCTGTGATGACCCAGGG 494
Db
347 AAGATCTGCCAAATCAGAACTTTAAATGCGAAGAGAGGTGACAGCAACACAG 406
Qy
495 AAGATCTGCCAAATCAGAACTTTAAATGCGAAGAGAGGTGACAGCAACACAG 554
Db
407 GTTTAAATGAGAGAGCTGAAACCAACAACTGTTTATCTAAGATATTTGACTTA 466
Qy
555 GTTTAAATGAGAGAGCTGAAACCAACAACTGTTTATCTAAGATATTTGACTTA 614
Db
467 AAAATATCAAAATAAATTTTTCAGCTTTCTCCAAAAA 509
Qy
615 AAAATATCAAAATAAATTTTTCAGCTTTCTCCAAAAA 657
Db

RESULT 3
AX921769
LOCUS 550 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 109 from Patent WO02068649.
ACCESSION AX921769
VERSION AX921769.1 GI:40215336
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS
JOURNAL
FEATURES
source
1. .550
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 89.9%; Score 457.8; DB 6; Length 550;
Best Local Similarity 96.8%; Pred. No. 28-106;
Matches 483; Conservative 0; Mismatches 2; Indels 14; Gaps 1;
Qy 25 GTTCCTGCTGCTGGAGTCTTTCTGCTCCCACTGAGAGCGAGCT-----GTG 70
Db 1 GTTCCTGCTGCTGGAGTCTTTCTGCTCCCACTGAGAGCGAGCTGTTCTGTTGCAGTG 60
Qy 71 TGAATATGATTTGGCGAGGAGATCAATATAGGCTTAGGCGGAGGAGAGTGTACCA 130
Db 61 TGAATATGATTTGGCGAGGAGATCAATATAGGCTTAGGCGGAGGAGAGTGTACCA 120
Qy 131 CCTCCTGAGTGTGGGCTATGCTGGAGCCGGTGTAGAGAGCTCAGCAAGAGGAA 190
Db 61 TGAATATGATTTGGCGAGGAGATCAATATAGGCTTAGGCGGAGGAGAGTGTACCA 120
Qy 131 CCTCCTGAGTGTGGGCTATGCTGGAGCCGGTGTAGAGAGCTCAGCAAGAGGAA 190
Db 121 CCTCCTGAGTGTGGGCTATGCTGGAGCCGGTGTAGAGAGCTCAGCAAGAGGAA 180
Qy 191 CCACCAACTGAAAGTCCGGATCTTCGACCTGGTCCAGAGAGAGAGAGATCAGGGTGCA 250
Db 181 CCACCAACTGAAAGTCCGGATCTTCGACCTGGTCCAGAGAGAGAGAGATCAGGGTGCA 240
Qy 251 GCTGAGACTCAAGTCCCTGAGCTGGAGCTGATCTCCAGAGAGCTCTCAGTCAAGAGCT 310
Db 241 GCTGAGACTCAAGTCCCTGAGCTGGAGCTGATCTCCAGAGAGCTCTCAGTCAAGAGCT 300
Qy 311 GGGGTGATGTGGAATGGTCTGTGATGACCCAGGGGAGAGTTCTGCCAAATCAGACAA 370
Db 301 GGGGTGATGTGGAATGGTCTGTGATGACCCAGGGGAGAGTTCTGCCAAATCAGACAA 360
Qy 371 TTTAAATGCCAGAGAGAGGTGACAGGCAACACAGGTTTAAATGAAGACAAGCTGAAAC 430
Db 481 GCTTTCTCCAAAAA 499

251 GCTGAGACTCAAGTCCCTGAGCTGGAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGAGCT 310
Db
241 GCTGAGACTCAAGTCCCTGAGCTGGAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGAGCT 300
Qy
311 GGGGTGATGTGGAATGGTCTGTGATGACCCAGGGGAGAGTTCTGCCAAATCAGACAA 370
Db
301 GGGGTGATGTGGAATGGTCTGTGATGACCCAGGGGAGAGTTCTGCCAAATCAGACAA 360
Qy
371 TTTAAATGCCAGAGAGAGGTGACAGGCAACACAGGTTTAAATGAAGACAAGCTGAAAC 430
Db
361 TTTAAATGCCAGAGAGAGGTGACAGGCAACACAGGTTTAAATGAAGACAAGCTGAAAC 420
Qy
431 AACACAAAATGTTTATCTAAGATATTTGACTTAAATATCAAAATCAAACTTTTGCA 490
Db
421 AACACAAAATGTTTATCTAAGATATTTGACTTAAATATCAAAATCAAACTTTTGCA 480
Qy
491 GCTTTCTCCAAAAA 509
Db
481 GCTTTCTCCAAAAA 499

RESULT 4
AX921771
LOCUS 494 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 111 from Patent WO02068649.
ACCESSION AX921771
VERSION AX921771.1 GI:40215337
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS
JOURNAL
FEATURES
source
1. .494
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 89.6%; Score 456; DB 6; Length 494;
Best Local Similarity 97.2%; Pred. No. 57e-106;
Matches 480; Conservative 0; Mismatches 0; Indels 14; Gaps 1;
Qy 25 GTTCCTGCTGCTGGAGTCTTTCTGCTCCCACTGAGAGCGAGCT-----GTG 70
Db 1 GTTCCTGCTGCTGGAGTCTTTCTGCTCCCACTGAGAGCGAGCTGTTCTGTTGCAGTG 60
Qy 71 TGAATATGATTTGGCGAGGAGATCAATATAGGCTTAGGCGGAGGAGAGTGTACCA 130
Db 61 TGAATATGATTTGGCGAGGAGATCAATATAGGCTTAGGCGGAGGAGAGTGTACCA 120
Qy 131 CCTCCTGAGTGTGGGCTATGCTGGAGCCGGTGTAGAGAGCTCAGCAAGAGGAA 190
Db 61 TGAATATGATTTGGCGAGGAGATCAATATAGGCTTAGGCGGAGGAGAGTGTACCA 120
Qy 131 CCTCCTGAGTGTGGGCTATGCTGGAGCCGGTGTAGAGAGCTCAGCAAGAGGAA 190
Db 121 CCTCCTGAGTGTGGGCTATGCTGGAGCCGGTGTAGAGAGCTCAGCAAGAGGAA 180
Qy 191 CCACCAACTGAAAGTCCGGATCTTCGACCTGGTCCAGAGAGAGAGAGATCAGGGTGCA 250
Db 181 CCACCAACTGAAAGTCCGGATCTTCGACCTGGTCCAGAGAGAGAGAGATCAGGGTGCA 240
Qy 251 GCTGAGACTCAAGTCCCTGAGCTGGAGCTGATCTCCAGAGAGCTCTCAGTCAAGAGCT 310
Db 241 GCTGAGACTCAAGTCCCTGAGCTGGAGCTGATCTCCAGAGAGCTCTCAGTCAAGAGCT 300
Qy 311 GGGGTGATGTGGAATGGTCTGTGATGACCCAGGGGAGAGTTCTGCCAAATCAGACAA 370
Db 301 GGGGTGATGTGGAATGGTCTGTGATGACCCAGGGGAGAGTTCTGCCAAATCAGACAA 360
Qy 371 TTTAAATGCCAGAGAGAGGTGACAGGCAACACAGGTTTAAATGAAGACAAGCTGAAAC 430
Db 481 GCTTTCTCCAAAAA 499


```

Db      361  TTTAAATCCAGAGGAGTGCACAGGCAACACAGGTTTAAATGACAGCAAGCTGAAC 420
Qy      431  AACACAAAACCTGTTTATCTAAGATATTGATCTTAAATAATCAAAATAAATCTTTGCA 490
Db      421  AACACAAAACCTGTTTATCTAAGATATTGATCTTAAATAATCAAAATAAATCTTTGCA 480
Qy      491  GCTTTCTCCAAAAA 504
Db      481  GCTTCTCCAAAAA 494

RESULT 5
HSA318881
LOCUS      HSA318881      493 bp      mRNA      linear      PRI 14-JAN-2002
DEFINITION Homo sapiens mRNA for XAGE-3 protein.
ACCESSION  AJ318881
VERSION     XAGE-3 gene.
KEYWORDS    Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
REFERENCE   1  Zendman,A.J.W., van Kraats,A.A., Weidle,U.H., Ruiter,D.R. and Van
AUTHORS     Muijen,G.N.P.
TITLE       Expression profile of members of the XAGE cancer/testis antigen
            family
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 493)
AUTHORS     Zendman,A.J.W.
TITLE       Direct Submission
JOURNAL     Submitted (14-AUG-2001) Zendman A.J.W., Department of Pathology,
            University Medical Center St Radboud, Geert Grooteplein Zuid 24, PO
            Box 9101 6500 HB Nijmegen, NETHERLANDS
FEATURES    Location/Qualifiers
            1..493
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /chromosome="X"
               /map="Xp11.21-Xp11.22"
               /tissue_type="Ewing sarcoma"
            1..493
               /gene="XAGE-3"
               /gene_start=1
               /product="XAGE-3 protein"
               /protein_id="CAC83008.1"
               /db_xref="GI:18157212"
               /db_xref="SWISS-PROT:Q8WTP9"
               /translation="MIWRGSRSTYRPRRSVPPPELIGPMLPEGPDEPQEEPPPTESR
               DPAPQQRREDQGAETQVPLEADLOELQSQKSGECNGPDQDKILPKSEQFQMP
               EGGDRQPV"

ORIGIN
Query Match      86.3%; Score 439.4; DB 9; Length 493;
Best Local Similarity 99.8%; Pred. No. 1e-101;
Matches 440; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      69  TGTGAATATGATTTGGGAGGAGATCAACATATAGCGCTAGCCGAGGAGAGAGTGATC 128
Db      53  TGTGAATATGATTTGGGAGGAGATCAACATATAGCGCTAGCCGAGGAGAGAGTGATC 112
Qy      129  CACCTCTGAGCTGATTTGGGCTTATGCTGGAGCCCGGTGATGAGGAGCTTCAGCAAGAG 188
Db      113  CACCTCTGAGCTGATTTGGGCTTATGCTGGAGCCCGGTGATGAGGAGCTTCAGCAAGAG 172
Qy      189  AACACCAACATGAAGTCGGATCTCTGACCTGTCAGGAGAGAGAGATCAGGGTG 248
Db      173  AACACCAACATGAAGTCGGATCTCTGACCTGTCAGGAGAGAGAGATCAGGGTG 232
Qy      249  CAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGA 308

```

```

Db      233  CAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGA 292
Qy      309  CTGGGGTGAATGTGAAATGGTCTCTGATGACACAGGGGAGGATTCTGCCAAATCAGAAC 368
Db      293  CTGGGGTGAATGTGAAATGGTCTCTGATGACACAGGGGAGGATTCTGCCAAATCAGAAC 352
Qy      369  AATTTAAATGCCAGAGAGGTGACAGGCAACCCACAGGTTTAAATGAAGACAGCTGAA 428
Db      353  AATTTAAATGCCAGAGAGGTGACAGGCAACCCACAGGTTTAAATGAAGACAGCTGAA 412
Qy      429  ACAACACAAACCTGTTTATCTAAGATATTGATCTTAAATAATCAAAATAAATCTTTG 488
Db      413  ACAACACAAACCTGTTTATCTAAGATATTGATCTTAAATAATCAAAATAAATCTTTG 472
Qy      489  CAGCTTTCTCCAAAAA 509
Db      473  CAGCTTTCTCCAAAAA 493

RESULT 6
HSA318880
LOCUS      HSA318880      620 bp      mRNA      linear      PRI 14-JAN-2002
DEFINITION Homo sapiens mRNA for XAGE-2 protein.
ACCESSION  AJ318880
VERSION     XAGE-2 gene.
KEYWORDS    Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
REFERENCE   1  Zendman,A.J.W., van Kraats,A.A., Weidle,U.H., Ruiter,D.R. and Van
AUTHORS     Muijen,G.N.P.
TITLE       Expression profile of members of the XAGE cancer/testis antigen
            family
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 620)
AUTHORS     Zendman,A.J.W.
TITLE       Direct Submission
JOURNAL     Submitted (14-AUG-2001) Zendman A.J.W., Department of Pathology,
            University Medical Center St Radboud, Geert Grooteplein Zuid 24, PO
            Box 9101 6500 HB Nijmegen, NETHERLANDS
FEATURES    Location/Qualifiers
            1..620
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /chromosome="X"
               /map="Xp11.21-Xp11.22"
               /tissue_type="Ewing sarcoma"
            1..620
               /gene="XAGE-2"
               /gene_start=1
               /product="XAGE-2 protein"
               /protein_id="CAC83007.1"
               /db_xref="GI:18157210"
               /db_xref="SWISS-PROT:Q96GT9"
               /translation="MSWRGSRSTYRPRRSVPPPELIGPMLPEGPDEPQEEPPPTKSR
               NPTPDOKREDQGAETQVPLEADLOELQKCTKTGDCGEGTDYKGIILPKAEHFKMP
               EAGEGKSV"

ORIGIN
Query Match      66.3%; Score 337.4; DB 9; Length 620;
Best Local Similarity 80.2%; Pred. No. 1.4e-75;
Matches 437; Conservative 0; Mismatches 71; Indels 37; Gaps 2;

Qy      1  GAGTTGTAGGGTGTGAGGTCGCGCTTCTGCTGCTCGACTTTTCTGTGCCACTTTTCTGTCCCACTGAGA 60
Db      77  GAGCTGTGAGGTGTGAGGGGACGCTTCAGCCGCTCTGACTCTTCTCTCTCTACTGAGA 136

```


182	GCAGGAGATCAACATATAGGCTAGGCCAAGAAGATTTTACAGCTCTGAGCTGAT	241
145	TGGGCTATGCTGGAGCCCGGTGATGAGGAGCCTCAGCAAGAGGAACCAACACTGAAG	204
242	TGGGGCTATGCTTGAACCACTGATGAAGACCTTAAAGAAGAAACCAACCACTAAAG	301
205	TCGGGATCCTGCACCTG--GTTCAGGAGAGAGAAGAATCAGGGTGCAGCTGAGACTCA	262
302	TCGGAACTCTACCTCTGACTCNAGAAGAGAGAAGATGATCAGGGTGCAGCTGAGATTC	361
263	GTGCTGTGACTGGAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGAGACTGGGGTGAATGT	322
362	GTGCTGTGACTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAGACTGGGATGATGT	421
323	GGAAATGGTCTCTGATGACGAGGGAAGATTCGCAAAATCAGACAATTTTAAATGCCA	382
422	GAAGGTGGTACTGATGTCAAGGGGAAGATTTCAACAAAGCAGAGCACTTTAAATGCCA	481
383	GAAGGAGGTGACAGGCCACACAGGTTTAAATGAGACAGCTTGAACACACACAAACTG	442
482	GAAGCAGGTGAAGGGAATTCACAGGTTTAAAGGAAGATTAAGCTGAAACACACAC-AA	540
443	TTTTTATCTTAAGATATTTGACTTAAAAATATCAAAATAAACTTTTGCAGCTTTTCTC	502
541	TTTTTATATTAGATATTTTACTTTTAAATAATCTTAAATAAAGTTTTAAGCTTTTCTC	600
503	AAAAAAA	509
601	AAAAAAA	607

RESULT 10
HSL185E6A/c
LOCUS
DEFINITION
HSL185E6A linear DNA 20587 bp
Human DNA sequence from cosmid L129H7, Huntington's Disease Region,
chromosome 4p16.3 contains Pseudogene and CpG island.
ACCESSION
Z68274.1 GI:1130686
VERSION
4P16.3; CpG island; pseudogene.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 20587)
AUTHORS
Buck,D.
TITLE
Direct Submission
JOURNAL
Submitted (12-DEC-1995) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, England. E-mail enquiries: humquerv@sanger.ac.uk

flow-sorted human chromosome 4 from a Hamster-Human hybrid cell line (UV20HL21-27) containing human chromosomes 4, 8 and 21. VECTOR: scos1

L195E6 is contained in a clone contig spanning 2MB which is described in Baxendale et al, Nature Genetics 4 (1993) 181-186. See also Myers et al, Cytogenet Cell Genet. 66 (1994) 218-230.

[illegible]

QY 121 AAGTGTACCACTCTGAGCTGATTTGGCCCTATGCTGGACCCGGTGATGAGAGCCTCA 180
Db 18093 GTATGTCCAGCTCTCTGAGTACTGGCCCTGTCTGTAGCCAGATGAGAGCCTCA 18034
QY 181 GCAAGAGAAACCACTAACTGAAAGTGGGATCTGTGACCTGGTGTGAGAGAGAGAAAGA 240
Db 18033 GCAAGAGAAACCACTAACTGAAAGTGGGATCTTT--ACCTGGCCAGAGAGAGAGAA 17976
QY 241 TCAGGGTGACGTGAGACTCAAGTGCTGACCTGGAGCTGATCTCCAGAGCTGTCTCA 300
Db 17975 TCAGGGTGCTGAGATTTCTGTGCTGTACAGAGAGCTGATCTCCGAGAGCTGTCTC- 17917
QY 301 GTCAAGACTGGGGGTGAATGTGGAATGTCTGTGATGACAGGGGAAATTTCTGCAAA 360
Db 17916 ----AAAGACTGGGATGAATGTGGAGATGCTCTGTATGTC--GGGGAATATTCTGCCGAA 17862
QY 361 ATCAGACCAATTTAAATGCTCAGAGAGAGGTGACAGCCACACAGGTTTAATGAGAC 420
Db 17861 ATCAGACCAATTTAAATGCTCAGAGAGAGGTGAGGACACACAGGTTTCCAGGAGAC 17802
QY 421 AAGCTGAAACCAACCAAACTGTTTTTATCTAAGATATTGACTTAAAA 469
Db 17801 AAGCTGAAACCAATGC-AAAGCTGTTTTATATTAGATAGTGTACTTAAAA 17754

RESULT 11
HS193G15/c
LOCUS Human DNA sequence from clone RP1-193G15 on chromosome 4, complete
DEFINITION
ACCESSION AL117391
VERSION AL117391.1 GI:5911814
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 62493)
McMurray, A.
Direct Submission
Submitted (05-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: En: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep RP1-193G15 is from the library RP1-1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pCYPAC2.

FEATURES
source

Location/Qualifiers
1..62493
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/clone="RP1-193G15"
/clone_lib="RP1-1"

ORIGIN

Query Match 57.8%; Score 294.2; DB 9; Length 62493;
Best Local Similarity 82.5%; Pred. No. 2.1e-64;
Matches 387; Conservative 0; Mismatches 73; Indels 9; Gaps 4;
QY 1 GAGTTGTGAGGGTGTGAGGGTTCGGTTCTCTGCTGTCTGTGAGCTTTTCTGTGCTCACTGAGA 60
Db 54304 GACCTGTGAGGGTGTGAGGGTTCGGTTCTCTGCTGTCTGTGAGCTTTTCTGTGAGCTCACTGAGA 54245
QY 61 CGAGCTGTGTGAATATGATTTGGCGAGGAAGATCAACATATAGGCTTAGGCTAGGCGGAGGAG 120
Db 54244 AGCAGCCCTTTGAAATATGAGTTGGCAAGGAAGATCAGCATGTAGGCTTAGGCTTAGGCGGAGG 54185
QY 121 AAGTGTACCACTCTCTGAGCTGATTGGGCTTATGCTGGAGCCCGGCTGATGAGAGGCTCTCA 180
Db 54184 GTAATGTGAGCTCTCTGAGCTTACTGGGCTGTGCTTGAGCCCGGCTGATGAGAGGCTCTCA 54125
QY 181 GCAAGAGAACCACTCAAACTGAAAGTCCGGATCTCTGACCTGTGTCAGGAGAGAGAGAGAA 240
Db 54124 GCAAGAGAACCACTCAAACTGAAAGTCCGGATCTT--ACCTGGCCAGGAGAGAGAGAGAA 54067
QY 241 TCAGGGTGACGTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGAGAGCTGTCTCA 300
Db 54066 TCAGGGTGCTGTGAGATTTCTGTGCTTGACAGGAGAGCTGATCTCCGAGAGCTGTCTC- 54008
QY 301 GTCAAGACTGGGGGTGAATGTGGAATGTCTGTGATGACAGGGGAGAGATTTGTCCTCAAA 360
Db 54007 ----AAAGACTGGGATGAATGTGAGATGCTCTGTATGTC--GGGGAATATTCTGCCGAA 53953
QY 361 ATCAGAACAAATTTAAATGTCAGAGAGAGGTGACAGGCAACCACTGATTTAAATGAAGAC 420
Db 53952 ATCAGAGCAATTTAAATGTCAGAGAGAGGTGAGGCGACACCACTGATTTCCAGGAGAGAC 53893
QY 421 AAGCTGAAACCAACCAAACTGTTTTTATCTAAGATATTGACTTAAAA 469
Db 53892 AAGCTGAAACCAATGC-AAAGCTGTTTTATATTAGATAGTGTACTTAAAA 53845

RESULT 12
HS193G15/c

LOCUS HS193G15 62493 bp DNA linear PRI 23-NOV-1999
DEFINITION Human DNA sequence from clone 193G15 on chromosome 4 Contains the 5' part of the RGS12 gene for regulator of G-protein signalling 12, an RPL7A (60S Ribosomal Protein 7A (SURF3)) pseudogene, ESTs, an STS, GSSs and a putative CpG island, complete sequence.
ACCESSION AL096838
VERSION AL096838.1 GI:5459253
KEYWORDS HTG; 60S Ribosomal Protein 7A; CpG Island; G-protein; RGS12; RPL7A; SURF3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 62493)
McMurray, A.
Direct Submission
Submitted (10-JUL-1999) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em., EMSL; Sw., SWISSPROT; Tr., TREMBL
IMPORTANT: This sequence is not the entire insert of clone 193G15. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.
The true left end of clone 1E6 is at 6371 in this sequence. The true right end of clone 361H4 is at 104 in this sequence. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.
This sequence was generated from part of bacterial clone contigs of human chromosome 4, constructed by the Sanger Centre Chromosome 4 Mapping Group.
193G15 is from the library RPC11 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/VECTOR:pcypac2>.

FEATURES

Location/Qualifiers
1..62493
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/clone="RP1-193G15"
/clone_lib="RPC1-1"

ORIGIN

Query Match 57.8%; Score 294.2; DB 9; Length 62493;
Best Local Similarity 82.5%; Pred. No. 2.1e-64;
Matches 387; Conservative 0; Mismatches 73; Indels 9; Gaps 4;
QY 1 GAGTTGTGAGGGTGTGAGGGTCCGCTTCCTGCTGTCTGGACTTTTCTGTCCTCCACTGAGA 60
Db 54304 GACCTGTGAGGGTGTGAGGGTTCGCTTCCTGGTCTGGACTTTTACCCACTACTGAGA 54245
QY 61 CGCAGCTGTGGAATATGATTTGGCGAGGAAGATCAACATATAGCGCTAGCCGAGGAG 120
Db 54244 AGCAGCGGTTTGAAATATGAGTGGCAAGGAAGATCAGCATGTAGGCTTAGACCAAGACG 54185
QY 121 AAGTGTACCACTCTGAGCTGATTTGGGCTATGCTGAGCCCGGTGATGAGGACCTCA 180
Db 54184 GTATGTGACGCTCTGAGCTAATCTGGGCTGTGCTTGAGCCCACTGATGAGCAGCTCA 54125
QY 181 GCAAGAGGAACCAACCACTGAAAGTGGGATCCTGCACTGTGTACGAGAGGAAGAAGA 240
Db 54124 GCAAGAGGAACCAACCACTGAAAGTGGGCTCTT--ACCTGGCCAGGAGAGGAAGAAGA 54067
QY 241 TCAGGTGCAGCTGAGACTCAAGTCTGCTGACCTGGAAGCTGATCTCCAGGAGCTGTCTCA 300
Db 54066 TCAGGTGTCTGTGAGATTTCTGTGCTTGACCAAGAGCTGATCTCCGAGAGCTGTCTC- 54008
QY 301 GTCAAGACTGGGGGTGAATGTGGAATGGTCTGATGACCAAGGGAAGATTTCTCCAAA 360
Db 54007 ----AAAGACTGGGATGAATGTGGAGATGGTCTGATGTCC--GGGGAATATTCTCCGAA 53953
QY 361 ATCAGACAAATTTAAATCCAGAGGGGTGACAGGCAACCAAGGTTTAAATCAAGAC 420
Db 53952 ATCAGAGCAATTTAAATTCAGAGAGGATGAAGGCAACCAAGGTTTCCAGAGAGAC 53893
QY 421 AAGCTGAAACAAACAAAACTGTTTTTATTAAGATATTGTACTTAAAA 469
Db 53892 AAGCTGAAACAAATGC-AAACTGTTTTTATTATAGATACGTGACTTAAAA 53845

RESULT 13
AL645949/c 145151 bp DNA linear PRI 23-OCT-2002
LOCUS
DEFINITION Human DNA sequence from clone RP11-357G3 on chromosome 4, complete

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AL645949 GI:24366471
HTG.
Homo sapiens (human)

REFERENCE

1 (bases 1 to 145151)

AUTHORS

Direct Submission

TITLE

Submitted (16-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,

JOURNAL

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk

COMMENT

On Oct 24, 2002 this sequence version replaced gi:24210295.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.: EMBL; Sw.: SWISSPROT; Tr.: TREMBL; Wp.: WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP11-357G3 is from the library RPC1-11.2 constructed by the group of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6. Location/Qualifiers

1..145151
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/clone="RP11-357G3"
/clone_lib="RPC1-11.2"

FEATURES

source

1..145151
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/clone="RP11-357G3"
/clone_lib="RPC1-11.2"

ORIGIN

Query Match 57.8%; Score 294.2; DB 9; Length 145151;
Best Local Similarity 82.5%; Pred. No. 2.2e-64;
Matches 387; Conservative 0; Mismatches 73; Indels 9; Gaps 4;
QY 1 GAGTTGTGAGGGTGTGAGGGTCCGCTTCCTGCTGTCTGGACTTTTCTGTCCTCCACTGAGA 60
Db 102337 GACCTGTGAGGGTGTGAGGGTTCGCTTCCTGGTCTGGACTTTTACCCACTACTGAGA 102278
QY 61 CGCAGCTGTGTGAATATGATTTGGCGAGGAAGATCAACATATAGGCTAGGCGGAGGAG 120
Db 102277 AGCAGCGGTTTGAAATATGATTTGGCAAGGAAGATCAGCATGTAGGCTTAGACCAAGACG 102218
QY 121 AAGTGTACCACTCTGAGCTGATTTGGGCTTATGCTGAGCCCGGTGATGAGGAGCTCA 180
Db 102217 GTATGTGAGGCTCTGAGCTAATCTGGGCTGTGCTTGGAGCTCTTTACCCACTACTGAGA 102158
QY 181 GCAAGAGGAACCAACCACTGAAAGTGGGATCCTGCACTGTGTCAGGAGAGGAAGAAGA 240
Db 102157 GCAAGAGGAACCAACCACTGAAAGTGGGCTCTT--ACCTGGCCAGGAGAGGAGAGA 102100

QY	241	TCAGGGTCGAGCTGAGACTCAAGTGCCTGACCTGGAGAGCTGATCTCCAGAGCTGTCTCA	300
Db	102099	TCACGGTCTGCTGAGATCTTGTGCTGTGACCGAAGAGCTGATCTCCGGAGCTGTCTC-	102041
QY	301	GTCAAGAGACTGGGGGTGAATGTGGAATATGGTCTCTGATGACGAGGGGAAGATTTCTGCCAA	360
Db	102040	----AAAGACTGGGATGAATGTGAGATGGTCTCTGATGCC-GGGGAATATTCTGCCGAA	101986
QY	361	ATCAGAACAAATTTAAATGCCAGAGAGGTGACAGCCACACAGGTTTAATGAGAC	420
Db	101985	ATCAGAGCAATTTAAATGTCCAGAGAGGAGTGAAGGCAACACAGGTCTCCAAAGGAGAC	101926
QY	421	AAAGCTGAACACACACAAACACTGTTTTATCTAAAGATATTGACTTAAAA	469
Db	101925	AAAGCTGAACAAATGC-AAACTGGTTTTATATTAGATACGTGACTTAAAA	101878
RESULT 14			
AX078298	AX078298	580 bp	DNA
LOCUS	Sequence 102 from Patent WO0107471.	linear	PAT 22-FEB-2001
DEFINITION	AX078298		
ACCESSION	AX078298.1	GI:13157989	
VERSION			
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 Hillman, J.L., Lal, P., Tang, Y.T., Yue, H., Au-Young, J., Bandman, O.,		
JOURNAL	Azinzai, Y., Yang, J., Lu, D.A., Baughn, M.R., Patterson, C. and Shah, P.		
	Cell cycle and proliferation proteins		
	Patent: WO 0107471-A 102 01-FEB-2001;		
FEATURES	Incyte Genomics, Inc. (US)		
source	Location/Qualifiers		
	1..580		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN	/note="Incyte ID No: 4764233CB1"		
Query Match	57.5%;	Score 292.6;	DB 6; Length 580;
Best Local Similarity	79.7%;	Fred. No. 4.1e-64;	
Matches	374; Conservative	0; Mismatches 59;	Indels 36; Gaps 1;
QY	1	GAGTTGTGAGGGTGTGAGGGTCCGGTTCCTGCTCTCTGGACTTTTCTGCCACTGAGA	60
Db	112	GAGCTGTGAGAGTGTGAGGGGCAGCTTCCAGCGCTGGAATCTTTCTCTCTACTGAGA	171
QY	61	CCGACGTGT-----	84
Db	172	CGCAGCTTATAGTTCGGCAGGCCAGCTCTCCAGGAAGTGAATATGATGATGTTG	231
QY	85	GCAGGAGAGATCAACATATAGCCCTAGCCCGAGGAGAAGTGTACCACTCTCTGAGCTGAT	144
Db	232	GCAGGAGAGATCAACATATAGCCCTAGCCCGAAGAAGAAGTTTACAGCCTCTGAGCTGAT	291
QY	145	TGGGGCTATGCTGGAGCCGGTGTGAGAGCCTCAGCAGGAGAACCCAACTGAAAG	204
Db	292	TGGGGCTATGCTGTAACCACTGATGAAGCCTTAAGAGAGGAACACCCTAAAG	351
QY	205	TGGGGATCTTCGACTTGTGTCAGGAGAGAGAAGATCATGGGTGCAGCTGAGACTCAAGT	264
Db	352	TGGGAATCTTACACCTGATCAGAGAGAGAAGATGATCAGGGTGCAGCTGAGATTCAAGT	411
QY	265	GCTTGACCTGGAGCTGATCTCCAGAGCTGTCTCAGTCAAGAGCTGGGGTGAATGTGG	324
Db	412	GCTTGACCTGGAGCGATCTCCAGAGCTATGTGACAGAAAGACTGGGATGATGTGA	471
QY	325	AAATGCTCTGTATGACCGAGGGAAGATTCTGCCAAATTCAGAACAAATTTAAATGCCAG	384
Db	472	AGTGTGATCTGATGTCAAGGGGAAGATTCTACCAAGAGCAGAGCACTTTAAATGCCAG	531

/tissue_type="Lung, carcinoma, large cell

undifferentiated"

/clone_lib="NIH_MGC_69"

/lab_host="DH10B"

/note="Vector: PCMV-SPORT6"

1. .529

gene

/gene="GAGED2"

/note="synonyms: XAGE1, XAGE-1"

/db_xref="LocusID:9503"

/db_xref="MIM:300289"

184. .429

/codon_start=1

/product="GAGED2 protein"

/protein_id="AAH09538.1"

/db_xref="GI:16306935"

/db_xref="LocusID:9503"

/translation="MESPKNKQOLKVLHLSRKKIRIQLRSQCATWKVICKSCI

SQTFGINLDLGSVKVXLIIPKEHCKPEAGEEPQV"

CDS

ORIGIN

Query Match	42.1%;	Score 214.4;	DB 9;	Length 529;
Best Local Similarity	80.6%;	Pred. No. 4.4e-44;		
Matches 287;	Conservative 0;	Mismatches 66;	Indels 3;	Gaps 3;
QY 155	CTGGAGCCCGGTGATGAGGAGCCTCAGCAAGAGGAACCACTGAAAGTCGGGATCCT	214		
Db 171	CAGGAGCCAGTAATGGAGAGCCCAAAAGAAAGAACCCAGAGCTGAAGTCGGGATCCT	230		
QY 215	GCACCTGGTCAGGAGAGAGAAGATCAGGGTGCAGCTGAGACTCAAGTCCTGACCTG	274		
Db 231	ACACCTGGGACGACACAGAGAGAGATCAGGATACAGCTGAGATCCAGTCGGGACATG	290		
QY 275	GAAGCTGATCTCCAGAGCTGTCTCAGTCAAGACTGGGGTGAATGTGGAAT-GGTCC	333		
Db 291	GAAGGTGATCTGCAAGAGCTGCATCAGTCAACACCCGGGATAAATCTGGATTTGGGTTT	350		
QY 334	TGATGACCAAGGGGAAGATTCTGCCAAATTCAGAACAAATTTAAATGCCAGAGGAGGTGA	393		
Db 351	CGGCTCAAGGTGAAGATAATACCTTAAGAGGAACACTGTAAATGCCAGAGCAGGTGA	410		
QY 394	CAGGCAACCAAGGTTTAAATGAAGACAGCTGAACCAACACAAACTGTTTTTATCTAA	453		
Db 411	AGAGCAACCAAGTAAATGAAGACAGCTGAACCAACGC-AAGCTGGTTTATATTA	469		
QY 454	GATATTGACTTAAATATCAAAATAAATTTTCAGCTTTCTCCAAAAA	509		
Db 470	GATATTGACTT-AAACTATCTCAATAAGTTTTCAGCTTTCACCAAAAA	524		

Search completed: April 25, 2004, 18:57:27
Job time : 1351.3 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 17, 2004, 17:33:54 ; Search time 1687.16 Seconds
(without alignments)
16221.632 Million cell updates/sec

Title: US-10-051-835-1
Perfect score: 669
Sequence: 1 cacaacgagcagcagcagctt.....tttttgcagtcgacgctc 669

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :		GenEmbl:
1:	gb_ba:	
2:	gb_hg:	
3:	gb_in:	
4:	gb_om:	
5:	gb_ov:	
6:	gb_pat:	
7:	gb_ph:	
8:	gb_pl:	
9:	gb_pri:	
10:	gb_ro:	
11:	gb_sts:	
12:	gb_sy:	
13:	gb_un:	
14:	gb_vi:	
15:	em_ba:	
16:	em_fun:	
17:	em_hum:	
18:	em_in:	
19:	em_mu:	
20:	em_om:	
21:	em_or:	
22:	em_ov:	
23:	em_pat:	
24:	em_ph:	
25:	em_pl:	
26:	em_ro:	
27:	em_sts:	
28:	em_un:	
29:	em_vi:	
30:	em_hg_hum:	
31:	em_hg_inv:	
32:	em_hg_other:	
33:	em_hg_mus:	
34:	em_hg_pln:	
35:	em_hg_rtd:	
36:	em_hg_nam:	
37:	em_hg_vrt:	
38:	em_sy:	
39:	em_htgo_hum:	
40:	em_htgo_mus:	
41:	em_htgo_other:	

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	578.6	86.5	642	6	AX359705	Sequence
2	564	84.3	648	9	BC009232	Homo sapi
3	560.6	83.8	620	9	HSX318880	Homo sapi
4	549	82.1	580	6	AX078298	Sequence
5	507.8	75.9	611	6	AX226501	Sequence
6	287.2	42.9	493	9	HSX318881	Homo sapi
7	245	36.6	20587	9	HS18566A	Z68274 Human DNA s
8	245	36.6	62493	9	HS193G15	AL117391 Human DNA
9	245	36.6	62493	9	HSJ193G15	AL096838 Human DNA
10	245	36.6	145151	9	AL645349	AL645949 Human DNA
11	216	32.3	2848	9	HSX318891	Homo sapi
12	216	32.3	94514	9	AL953874	Human DNA
13	216	32.3	121324	9	BX293536	Human DNA
14	215	32.1	320	6	AX198429	Sequence
15	214	32.0	214	6	AX208978	Sequence
16	204	30.5	399	6	AR272357	Sequence
17	204	30.5	399	6	AR275338	Sequence
18	204	30.5	399	6	AX062442	Sequence
19	204	30.5	399	6	AX367359	Sequence
20	204	30.5	457	6	AR273057	Sequence
21	204	30.5	457	6	AR276638	Sequence
22	204	30.5	457	6	AX063163	Sequence
23	204	30.5	457	6	AX368080	Sequence
24	204	30.5	463	6	AX370580	Sequence
25	204	30.5	463	9	HSX290447	Homo sapi
26	204	30.5	529	9	BC009538	Homo sapi
27	204	30.5	626	9	AF251237	Homo sapi
28	204	30.5	637	6	AX455517	Sequence
29	204	30.5	740	9	HSX318878	Homo sapi
30	204	30.5	767	6	BD109661	BD109661 EST and e
31	204	30.5	781	6	AR273067	Sequence
32	204	30.5	781	6	AR276648	Sequence
33	204	30.5	781	6	AX063181	Sequence
34	204	30.5	781	6	AX368098	Sequence
35	182.8	27.3	528	6	AR275665	Sequence
36	182.8	27.3	528	9	AF055473	Homo sapi
37	181.2	27.1	530	9	HSU19143	Human GAGE-
38	181.2	27.1	538	6	AR028488	Sequence
39	181.2	27.1	538	6	BD132467	Isolated,
40	181.2	27.1	538	6	IS5851	Sequence 14
41	181.2	27.1	551	9	BC018052	Homo sapi
42	180	26.9	479	6	AR272358	Sequence
43	180	26.9	479	6	AR275939	Sequence
44	180	26.9	479	6	AX062443	Sequence
45	180	26.9	479	6	AX367360	Sequence
46	179.6	26.8	1245	6	AX285022	Sequence
47	178	26.6	461	6	AR273058	Sequence
48	178	26.6	461	6	AR276639	Sequence
49	178	26.6	461	6	AX063165	Sequence
50	178	26.6	461	6	AX368082	Sequence
51	178	26.6	481	9	HSX318879	Homo sapi
52	169.8	25.4	524	9	HSU19146	Human GAGE-
53	169.8	25.4	524	9	HSU19147	Human GAGE-
54	169.8	25.4	528	6	AX334151	Sequence
55	169.8	25.4	528	9	HSU19145	Human GAGE-
56	169.8	25.4	532	6	AR028491	Sequence
57	169.8	25.4	532	6	BD132470	Isolated,
58	169.8	25.4	532	6	IS5854	Sequence 17
59	169.8	25.4	539	6	AR028492	Sequence
60	169.8	25.4	539	6	BD132471	Isolated,
61	169.8	25.4	539	6	IS5855	Sequence 18
62	169.8	25.4	540	6	AR028490	Sequence
63	169.8	25.4	540	6	BD132469	Isolated,
64	169.8	25.4	540	6	IS5853	Sequence 16
65	169.8	25.4	637	9	BC024914	Homo sapi

cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kettman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 25 Row: m Column: 9
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.

FEATURES

Location/Qualifiers
1..648
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:16513 IMAGE:3960352"
/tissue_type="Placenta, choriocarcinoma"
/clone_lib="NIH_MGC_21"
/lab_host="DH10B-R"
/note="Vector: pOT87"
203..538
/codon_start=1
/product="Similar to G antigen 8"
/protein_id="AAH09232.1"
/db_xref="GI:14328032"
/translations="MSWGRSTYRPRRSLOPPELGAMLEPTDEPKBEKPTKSR
NPTPDQKREDQGAABIQVPLEADLQELCQTKTGDCGGTGVKGIKLPKAEHFMP
EAGEGKSOV"
EAGEGKSOV

BASE COUNT 207 a 134 c 162 g 145 t

ORIGIN

Query Match 84.3%; Score 564; DB 9; Length 648;
Best Local Similarity 95.8%; Pred. No. 6.2e-146;
Matches 598; Conservative 0; Mismatches 5; Indels 21; Gaps 1;
Qy 31 GTTCCTTGGACACCTGCTCAGTGTGCGATGTTCACTGGGCGATCTTCCCTCGACCCCTTT 90
Db 10 GTTCCTTGGACACCTGCTCAGTGTGCGATGTTCACTGGGCGATCTTCCCTCGACCCCTTT 69
Qy 91 GCCACGTGTGTACCGCTGGGGAGCTGTGAGTGTGAGGGGACGTTCCAGCCGCTCTGG 150
Db 70 GCCACGTGTGTACCGCTGGGGAGCTGTGAGTGTGAGGGGACGTTCCAGCCGCTCTGG 129
Qy 151 ACTCTTTCTCTCTACTGAGCGCAGCTATAGTTCGCGAGGCGAGTCTCCGCCAGGAAC 210
Db 130 ACTCTTTCTCTCTACTGAGCGCAGCTATAGTTCGCGAGGCGAGTCTCCGCCAGGAAC 189
Qy 211 GAATAGTGAATATAGTTGGCGAGGAAGATCAACATATAGGCTTAGGCGCAAGAAG 270
Db 190 GAATAGTGAATATAGTTGGCGAGGAAGATCAACATATAGGCTTAGGCGCAAGAAG 249
Qy 271 TTTCAGGCTCTGAGCTGATTGGGGCTATGCTTACTGGCTCCCTTTGTCCCGAGGAAC 330
Db 250 TTTCAGGCTCTGAGCTGATTGGGGCTATGCTTACTGGCTCCCTTTGTCCCGAGGAAC 288
Qy 331 CACTGATGAAGAGCTTAAGAGAGAAACCAACCCACTAAAGTCGGAATCTTACCTGA 390
Db 289 CACTGATGAAGAGCTTAAGAGAGAAACCAACCCACTAAAGTCGGAATCTTACCTGA 348
Qy 391 TCAGAGAGAGAGATGATCGGGTGCAGCTGAGATTCAGTGCCTGACCTGGAGCCGA 450
Db 349 TCAGAGAGAGAGATGATCGGGTGCAGCTGAGATTCAGTGCCTGACCTGGAGCCGA 408
Qy 451 TCTCAGGAGCTATGTCAGACAAAGACTGGGGATGGATGTGAAGTGTGCTACTGATGCAA 510
Db 409 TCTCAGGAGCTATGTCAGACAAAGACTGGGGATGGATGTGAAGTGTGCTACTGATGCAA 468
Qy 511 GGGGAAGATTCTACAAAGAGAGCACTTTAAATGCCAGACGAGTGAAGGGAATC 570

Db 469 GGGGAAGATTCTACAAAGAGAGCACTTTAAATGCCAGAGCGTGAAGGGAATC 528
Qy 571 ACAGGTTTAAAGGAAGTAAGCTGAAACACACAAACATGTTTTATATTAGATTTTAC 630
Db 529 ACAGGTTTAAAGGAAGTAAGCTGAAACACACAAACATGTTTTATATTAGATTTTAC 588
Qy 631 TTTAAAGAGCTCTTAATAAATTTT 654
Db 589 TTTAAATAATCTTAATAAGTTTT 612

RESULT 3

HS318880
LOCUS HS318880
DEFINITION Homo sapiens mRNA for XAGE-2 protein.
ACCESSION AJ318880.1 GI:18157209
VERSION XAGE-2 gene.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 Zandman,A.J.W., van Kraats,A.A., Weidle,U.H., Ruiter,D.R. and Van
Muijen,G.N.P.
AUTHORS Expression profile of members of the XAGE cancer/testis antigen
family

TITLE Unpublished
JOURNAL 2 (bases 1 to 620)
REFERENCE Zandman,A.J.W.
AUTHORS Direct Submission
TITLE Submitted (14-AUG-2001) Zandman A.J.W., Department of Pathology,
JOURNAL University Medical Center St Radboud, Geert Grooteplein Zuid 24, PO
Box 9101 6500 HB Nijmegen, NETHERLANDS

FEATURES

Source
1..620
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp11.21-Xp11.22"
/tissue_type="Ewing sarcoma"
gene
1..620
/gene="XAGE-2"
CDS
189..524
/gene="XAGE-2"
/codon_start=1
/product="XAGE-2 protein"
/protein_id="CAC83007.1"
/db_xref="GI:18157210"
/translation="MSWGRSTYRPRRSLOPPELGAMLEPTDEPKBEKPTKSR
NPTPDQKREDQGAABIQVPLEADLQELCQTKTGDCGGTGVKGIKLPKAEHFMP
EAGEGKSOV"

BASE COUNT 191 a 131 c 156 g 142 t

ORIGIN

Query Match 83.8%; Score 560.6; DB 9; Length 620;
Best Local Similarity 96.0%; Pred. No. 5.5e-145;
Matches 594; Conservative 0; Mismatches 4; Indels 21; Gaps 1;
Qy 36 TTGGACACCTGCTCAGTGTGCGATGTTCACTGGGCGATCTTCCCTCGACCCCTTTGCCCA 95
Db 1 TTGGACACCTGCTCAGTGTGCGATGTTCACTGGGCGATCTTCCCTCGACCCCTTTGCCCA 60
Qy 96 CTTGGTGACCGCTGGGAGCTGTGAGTGTGAGGGGACGTTCCAGCCGCTGAGACTCT 155
Db 61 CGTGGTGACCGCTGGGAGCTGTGAGTGTGAGGGGACGTTCCAGCCGCTGAGACTCT 120
Qy 156 TTCTCTCTCTACTGAGACGCGAGCTTAGTTCGCGAGGCGAGTCTCCCGAGGAACCTGAAT 215
Db 121 TTCTCTCTCTACTGAGACGCGAGCTTAGTTCGCGAGGCGAGTCTCCCGAGGAACCTGAAT 180
Qy 216 AGTGAATATAGTTGGCGAGGAAGATCAACATATAGGCTTAGGCGCAAGAAGTTTAC 275

181 AGTGAATATGAGTTGGCAGGAAGATCAATATAGGCTAGCCAGGAAGAGTTTAC 240
276 AGCTCCTGAGTGATTTGGGCTATGCTTACTGCTCCCTTTGTCCAGGAACCCACTG 335
241 AGCTCCTGAGTGATTTGGGCTATGCTT-----GAAACCCACTG 279
336 ATGAAGAGCTTAAAGAGGAAGAACCAACCACTTAAAGTCGGAATCTACACCTGATCAGA 395
280 ATGAAGAGCTTAAAGAGGAAGAACCAACCACTTAAAGTCGGAATCTACACCTGATCAGA 339
396 AGAGAGAGATGATCAGGCTGAGCTGAGATTCAGATTCAGTCCCTGACCTGGAACCCGATCTCC 455
340 AGAGAGAGATGATCAGGCTGAGCTGAGATTCAGATTCAGTCCCTGACCTGGAACCCGATCTCC 399
456 AGGAGCTATGTGACAGCAAGACTGGGATGGAATGTGAAGGTGGTACTGATGTCAAGGGGA 515
400 AGGAGCTATGTGACAGCAAGACTGGGATGGAATGTGAAGGTGGTACTGATGTCAAGGGGA 459
516 AGATTCTACCAAGAGAGAGCTTTAAATGCCAGAGCAGGTGAAGGGAAATCAGAGG 575
460 AGATTCTACCAAGAGAGAGCTTTAAATGCCAGAGCAGGTGAAGGGAAATCAGAGG 519
576 TTTAAAGGAAGATAGCTGAAACACACAACTGTTTTATATAGATATTTTACTTTAA 635
520 TTTAAAGGAAGATAGCTGAAACACACAACTGTTTTATATAGATATTTTACTTTAA 579
636 AGAGTCTTAATAAATTTT 654
580 AATATCTTAATAAAGTTTT 598

RESULT 4
AX078298 AX078298 580 bp DNA linear PAT 22-FEB-2001
LOCUS
DEFINITION Sequence 102 from Patent WO0107471.
ACCESSION AX078298
VERSION AX078298.1 GI:13157989
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Hillman,J.L., Lal,P., Tang,Y.T., Yue,H., Au-Young,J., Bandman,O.,
Azimzai,Y., Yang,J., Lu,D.A., Baughn,M.R., Patterson,C. and Shah,P.
TITLE Cell cycle and proliferation proteins
JOURNAL Patent: WO 0107471-A 102 01-FEB-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1..580
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 4764233CB1"
BASE COUNT 167 a 135 c 160 g 118 t
ORIGIN

Query Match 82.1%; Score 549; DB 6; Length 580;
Best Local Similarity 96.5%; Pred. No. 9.3e-142;
Matches 580; Conservative 0; Mismatches 21; Indels 1; Gaps 1;
QY 1 CACAAGCAGGACCGACTTCAGTGTGCAATGTTCCCTTGACACCTCCCTCAGTGTGCA 60
DB 1 CACAAGCAGGACCGACTTCAGTGTGCAATGTTCCCTTGACACCTCCCTCAGTGTGCA 60
QY 61 TTCACTGGGCACTTCTCCCTTCGACCCCTTTGCCACGTGTGACCCGCTGGGAGCTGTGA 120
DB 61 TTCACTGGGCACTTCTCCCTTCGACCCCTTTGCCACGTGTGACCCGCTGGGAGCTGTGA 120
QY 121 GAGTGTGAGGGGACGTTCCAGCCGCTGACCTTTCTCTCTACTGAGCCGACCTTA 180
DB 121 GAGTGTGAGGGGACGTTCCAGCCGCTGACCTTTCTCTCTACTGAGCCGACCTTA 180

QY 181 TAGGTCGAGGCCAGTCTCTCCAGGAATGAAATAGTGAATATAGTTCGCGAGGAG 240
DB 181 TAGGTCGAGGCCAGTCTCTCCAGGAATGAAATAGTGAATATAGTTCGCGAGGAG 240
QY 241 ATCAACATATAGGCTTAGGCTAGGAGAGTTTACAGCTCTCTGAGCTGATTTGGGGCTAT 300
DB 241 ATCAACATATAGGCTTAGGCTAGGAGAGTTTACAGCTCTCTGAGCTGATTTGGGGCTAT 300
QY 301 GCTTACTGGCTCCCTTTGTCTCCAGGAACCCACTGATGAAGAGCCCTAAAGAGAGAAACC 360
DB 301 GCTT-----GACCCACTGATGAAGAGCCCTAAAGAGAGAAACC 339
361 ACCCACTAAAGTCGAATCTACCTGATCAGAGAGAGAGATGATCAGGTCGAGC 420
340 ACCCACTAAAGTCGAATCTACCTGATCAGAGAGAGAGATGATCAGGTCGAGC 399
421 TGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTTCAGCAAAAGACTGG 480
400 TGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTTCAGCAAAAGACTGG 459
481 GATGATGTGAAGTGGTACTGATGTCAAGGGGAAGATTTACCAAAAGCAGGACACTT 540
460 GATGATGTGAAGTGGTACTGATGTCAAGGGGAAGATTTACCAAAAGCAGGACACTT 519
541 TAAAAATGCCAGAGCAGGTGAAGGGAAATCAGAGGTTTAAAGGAAGATAGCTGAAACAA 600
520 TAAAAATGCCAGAGCAGGTGAAGGGAAATCAGAGGTTTAAAGGAAGATAGCTGAAACAA 579
601 C 601
DB 580 C 580

RESULT 5
AX226501 AX226501 611 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION Sequence 7 from Patent WO0161009.
ACCESSION AX226501
VERSION AX226501.1 GI:15555699
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Malyankar,U.M., Tchernev,V.T., Padigaru,M., Taupier,R.J.,
Spytek,K.A., Majumder,K., Guo,X., Spaderna,S.K. and Boldog,F.L.
TITLE Polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 0161009-A 7 23-AUG-2001;
Curagen Corporation (US)
FEATURES
Location/Qualifiers
1..611
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 196 a 127 c 152 g 135 t
ORIGIN

Query Match 75.9%; Score 507.8; DB 6; Length 611;
Best Local Similarity 94.6%; Pred. No. 2.8e-130;
Matches 566; Conservative 0; Mismatches 8; Indels 24; Gaps 3;
QY 60 GTTCACTGGGCACTTCTCCCTTCGACCCCTTTGCCACGTGTGACCCCT-GGGGAGCTGT 118
DB 9 GTCACTGGGCACTTCTCCCTTCGACCCCTTTGCCACGTGTGACCCCTGGGGAGCTGT 68
QY 119 GAGAGTGTGAGGGCAGCTTCCAGCCGCTGGACTTTCTCTCTACTGAGAGCGAGCC 178
DB 69 GAGAGTGTGAGGGCAGCTTCCAGCCGCTGGACTTTCTCTCTACTGAGAGCGAGCC 128
QY 179 TATAGTTCGAGCCGAGCTCTCCAGGAATGAAATAGTGAATATAGTTCGCGAGGA 238

```
Db 129 TATAGTCCGAGCCAGTCTCCAGCAAGTGAATAGTGAATATGATGCTGGCGAGGA 188
Qy 239 AGATCAACATATAGGCTAGCCCAAGAGAGTTTACAGCTCTCTGAGCTGATGGGGCT 298
Db 189 AGATCAACATATAGGCTAGCCCAAGAGAGTTTACAGCTCTCTGAGCTGATGGGGCT 248
Qy 299 ATGCTTACTGGCTCCCTTTGTCACAGAACCCACTGATGAAGAGCCCTAAAGAGAGAAA 358
Db 249 ATGCTT-----GAACCCACTGATGAGAGCCCTAAAGAGAGAAA 287
Qy 359 CCACCCCTAAAGTCGGAATCTTACACCTGA--TCAGAGAGAGAGAGATGATCAGGGTG 416
Db 288 CCACCCCTAAAGTCGGAATCTTACACCTGACTCNAGAGAGAGAGATGATCAGGGTG 347
Qy 417 CAGCTGAGATTCAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGA 476
Db 348 CAGCTGAGATTCAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGA 407
Qy 477 CTGGGATGGATGGAAGTGGTACTGATGATCAAGGGAGAGATTCACCAAAAGCAGAGC 536
Db 408 CTGGGATGGATGGAAGTGGTACTGATGATCAAGGGAGAGATTCACCAAAAGCAGAGC 467
Qy 537 ACTTTAAATCCAGACAGCTGCAAGGGAATCACAGGTTTAAAGGAAGTAAAGTAAAGTGA 596
Db 468 ACTTTAAATCCAGACAGCTGCAAGGGAATCACAGGTTTAAAGGAAGTAAAGTAAAGTGA 527
Qy 597 ACAACACAACTGTTTTTATATTAGATATTTTACTTTAAAGAGTCTTAATAAATTTT 654
Db 528 ACAACACAACTGTTTTTATATTAGATATTTTACTTTAAAGATCTTAATAAAGTTT 585

RESULT 6
HSA318881 HSA318881 493 bp mRNA linear PRI 14-JAN-2002
LOCUS Homo sapiens mRNA for XAGE-3 protein.
DEFINITION
ACCESSION AJ318881
VERSION AJ318881.1 GI:18157211
KEYWORDS XAGE-3 gene.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Zendaan,A.J.W., van Kraats,A.A., Weidle,U.H., Ruiter,D.R. and Van
Muijen,G.N.P.
TITLE Expression profile of members of the XAGE cancer/testis-antigen
family
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 493)
AUTHORS Zendaan,A.J.W.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2001) Zendaan A.J.W., Department of Pathology,
University Medical Center St Radboud, Geert Grooteplein Zuid 24, PO
Box 9101 6500 HB Nijmegen, NETHERLANDS
FEATURES
Location/Qualifiers
1..493
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp11.21-Xp11.22"
/tissue_type="Ewing sarcoma"
1..493
/gene="XAGE-3"
61..396
/gene="XAGE-3"
/codon_start=1
/product="XAGE-3 protein"
/protein_id="CAC83008.1"
/db_xref="GI:18157212"
/translation="MIVRGSTYRPRRSVPPPLICPMLEPGDEBPQOEPPTESR
DPAQOEEREPQGAETQVPDLADLQELSKTGGCGNGPDQDGKILKSPQKWP
EGGDRQPV"
```

```
BASE COUNT 167 a 98 c 125 g 103 t
ORIGIN
Query Match 42.9%; Score 287.2; DB 9; Length 493;
Best Local Similarity 82.0%; Pred. No. 7.9e-69;
Matches 364; Conservative 0; Mismatches 58; Indels 22; Gaps 2;
Qy 213 ANTACTGAAATATGAGTTGGCGAGGAAGATCAACATATAGCCCTAGGCCACAGAGAACTT 272
Db 50 AAATGTGAAATATGAGTTGGCGAGGAAGATCAACATATAGCCCTAGGCCAGGAGAACTG 109
Qy 273 TACAGCCTCTGAGCTGATTTGGGGCTATGCTTACTTGGCTCCCTTTGTGCCAGGAACCA 332
Db 110 TACCACCTCTGAGCTGATTTGGGGCTATGCT-----GGAGCCCG 148
Qy 333 CTGATGAGAGAGCTAAAGAGGAACCAACCACTAAAGTCGGAATCCTACACCTGATC 392
Db 149 GTGATGAGAGAGCTCAGCAAGAGGAACCAACCACTGAAAGTCGGGATCCTGACCTGTC 208
Qy 393 AGAAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGAGTGCCTGGAAGCCGATC 452
Db 209 AGGAGAGAGAGAGATCAGGGTGCAGCTGAGACTCAAGTGCCTGGAAGCTGATC 268
Qy 453 TCAGAGAGCTATGTCAGACAAAGACTGGGATCGGATGGAAGTGGTACTGATGTCAGG 512
Db 269 TCCAGAGAGCTGCTCAGTCAAGAGACTGGGGGTGAATGTGGAATGGTCTGATGACCCAG 328
Qy 513 GGAAGATTTCAACAAAAGCAGAGACACTTTAAATGCCAGAGCAGGTGAAGGAAATCAC 572
Db 329 GGAGATTTCTGCAAAATCAGAACAAATTTAAATGCCAGAGGAGGTGACAGGCAACAC 388
Qy 573 AGGTTTAAAGAGAGATAGCTGAAACACAC-AAAAGTCTTTTATATTAGATATTTTACT 631
Db 389 AGGTTTAAATGAGACAAAGCTGAAACACACAAAAGTCTTTTATCTAAGATATTTGACT 448
Qy 632 TTAAGAGAGTCTTAATAAATTTTGG 655
Db 449 TAAATATATCGAATTAACCTTTTG 472

RESULT 7
HSL18556A/HSL18556A/c 20587 bp DNA linear PRI 23-NOV-1999
LOCUS Human DNA sequence from cosmid L129H7, Huntington's Disease Region,
chromosome 4p16.3 contains Pseudogene and CpG island.
DEFINITION
ACCESSION 268274.1 GI:1130686
VERSION 268274.1
KEYWORDS 4p16.3; CpG island; pseudogene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 20587)
AUTHORS Buck,D.
TITLE Direct Submission
JOURNAL Submitted (12-DEC-1995) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, England. E-mail enquiries: humquery@sanger.ac.uk
COMMENT
IMPORTANT:
This sequence is not the entire insert of clone L18556. It may be
shorter because we only sequence overlapping sections once, or
longer because we arrange for a small overlap between neighbouring
submissions. In addition this sequence has been finished according
to sequence map criteria as follows. An attempt is made to resolve
all sequencing problems, such as compressions and repeats, but not
necessarily within known annotated human repeat sequence elements
(e.g. Alu). Where unambiguous sequence has not been obtained,
there is an annotation using the 'unsure' feature key. Where the
size of a problem region is known only approximately the sequence
contigs on either side will be submitted as separate entries, with
annotation to indicate the order and positions of neighbours, with
what is known about the size and nature of the gap. Contigs have
entryname suffixes (A, B, ...) indicating the order of the contigs
with respect to the clone.
```

The sequence from clone L185B6 has been finished in more than one contig. This sequence (L185E6A) is separated from the following one (L185B6B) by a 900bp gap based on restriction data. 55bp VNTR made closure of gap impossible.

The true right end of clone L185B6 is at 10842 in this sequence. The true left end of clone L185B6 is at 1 in this sequence. L185B6 is from cosmid library LA04NC01 constructed at the Human Genome Center, Los Alamos National Laboratory, NM 87545 under the auspices of the U.S. Department of Energy. The library was constructed using flow-sorted human chromosome 4 from a Hamster-Human hybrid cell line (UV20HL21-27) containing human chromosomes 4, 8 and 21.

VECTOR: scos1

L185B6 is contained in a clone contig spanning 2mb which is described in Baxendale et al, Nature Genetics 4 (1993) 181- 186. See also Myers et al, Cytogenet Cell Genet. 66 (1994) 218-230.

FEATURES

```

source
    1..20587
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="4"
    /map="4p16.3"
    /clone="LA04NC01-185B6"
    /cell_line="UV20HL21-27"
    /clone_lib="LA04NC01"
    /complement(1370..1739)
    /note="match: STS L00885"
    2380..4066
    /note="L1 element fragment"
    repeat_region
    4115..4400
    /note="Alu repeat: matches 1..308 of consensus"
    repeat_region
    4427..4736
    /note="L1 element fragment"
    repeat_region
    4778..5099
    /note="L1 element fragment"
    repeat_region
    5203..5333
    /note="MSTC element fragment"
    repeat_region
    5632..5721
    /note="L1 element fragment"
    repeat_region
    5980..6714
    /note="L1 element fragment"
    CDS
    636072
    /note="Matches Trk-2h oncogene - 41aa of human RL7A_HUMAN"
    /pseudo
    /codon_start=1
    repeat_region
    6747..6778
    /note="15 copies of 2 mer 100 % conserved"
    repeat_region
    10906..11092
    /partial
    /note="Alu repeat: matches 1..191 of consensus"
    repeat_region
    11083..11214
    /partial
    /note="Alu repeat: matches 175..308 of consensus"
    repeat_region
    11352..11506
    /partial
    /note="Alu repeat: matches 308..151 of consensus"
    misc_feature
    13249..13799
    /note="putative CpG island"
    repeat_region
    15946..16238
    /note="Alu repeat: matches 1..308 of consensus"
    repeat_region
    17069..17143
    /note="MLT1D element fragment"
    repeat_region
    17589..17689
    /partial
    /note="Alu repeat: matches 308..210 of consensus"
    repeat_region
    17695..17739
    /partial
    /note="Alu repeat: matches 170..114 of consensus"
    repeat_region
    18268..18369
    /partial
    /note="Alu repeat: matches 155..38 of consensus"
    repeat_region
    19668..19742
    /note="MIR element fragment"

```

```

repeat_region 20313..20524
BASE COUNT 5439 a 5367 c 4601 g 5180 t
ORIGIN
Query Match 36.6%; Score 245; DB 9; Length 20587;
Best Local Similarity 70.9%; Pred. No. 6e-57;
Matches 426; Conservative 0; Mismatches 110; Indels 65; Gaps 5;
QY 58 ATGTTCACTGGGCGATCTCCCTTCGACCCCTTTGCCACGCTGTCACCGCTGGGAGCTG 117
Db 18267 AGTTTCACCTGGGCTGCTTCTCTTGGCCCTTTGCCACGCTGTCGACCTG 18208
QY 118 TGAGATGTGAGGGGACGCTTCAGCCGCTTCGACCTCTTCTCTCTCTACTGAGCGCAGC 177
Db 18207 TGAGGGGTGTGAGGGTTTCGCTTCTCTGCTTCTTACCCACTTCTGAGAGCAGC 18148
QY 178 CTATAGTCTCGGAGCCAGTCTCTCCGAGGAAGTGAATAGTGAATATGAGTTGGCGAG 237
Db 18147 CGTT-----TGAAATATGAGTTGGCAAG 18124
QY 238 AAGATCAACATATAGCGCTAGGCCAAGAAGAAGTTTACAGCCTCTCTGAGCTGATTCGGGC 297
Db 18123 AAGATCAGCATGTAGGCTTTAGACCAAGACGCTATGTGCAGCCTCTCTGAGCTAACCTGGGC 18064
QY 298 TATGCTTACTGGCTCCCTTTGTCCCGAGGAAGCCACTGATGAAGAGCCTTAAAGAGAGAA 357
Db 18063 TGTGCTT-----GAGCCCACTGATGAGCAGCCTCAGCAAGAGGA 18025
QY 358 ACCACCCACTAAAGTCGGAATCTTACCTGATCAGAGAGAGAGAGATGATCAGGGTGC 417
Db 18024 ACCACCACTTAAAGTCGGGGTCTT--ACCTGGCCAGGAGAGAGAGATGATCAGGGTGC 17967
QY 418 AGCTGAGATTTCAAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGAC 477
Db 17966 TGTGAGATTTCTGTGCTTGACCAAGAGCTGATCTCCGGAGCTCTCTCAAGA-----17912
QY 478 TGGGATGTGATGAAGTGGTACTGATGTCAAGGGGAGAGATTCTACCAAAAGCAGAGCA 537
Db 17911 CTGGGATGAATGTGGAGATGGTCTCTGATGTC-CGGGGAATATTCTCCGGAATACAGAGCA 17853
QY 538 CTTTAAATCCAGAGAGCAGTGAAGGGAATCAGAGGTTTAAAGGAAGATAGCTGAA 597
Db 17852 ATTAAATGCCAGAGAGGAGTGAAGGGAATCCAGAGTTTCCAGAGAGAGAGAGCTGAA 17793
QY 598 CAACACAACTGTTTTATATTAGATATTTTACTTTTAAAGAGTCTTAAATAATTTTGGC 657
Db 17792 CAATGCAAACTGGTTTTATATTATAGTACGTGACTTAAATATCTCAATACAGTTTCTCC 17733
QY 658 A 658
Db 17732 A 17732

```

RESULT 8

```

LOCUS HS193G15 62493 bp DNA linear PRI 05-MAR-2003
DEFINITION Human DNA sequence from clone RP1-193G15 on chromosome 4, complete
sequence.
ACCESSION AL117391
VERSION AL117391.1 GI:5911814
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 62493)
Direct Submission
AUTHORS McMurray,A.
TITLE Submitted (05-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humqy@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
----- Genome Center
COMMENT

```

Db	54002	CTGGGATGAATGTGGAGATGGTCTCTGATGTC-CGGGGAAATATTCTGCCGAAATCAGAGCA	53944
Qy	538	CTTTAAATATGCCAGAACGAGGTGAAGGGAAATACACAGGTTTAAAGGAAGATAAGCTGAA	597
Db	53943	ATTTAATTTGCCAAGAGGAGGTGAAGGGCAACCAACAGTTTCAAGGAAGACAGCTGAA	53884
Qy	598	CAACACAAACTGTTTATATTAGATATTTTACTTTAAAGAGCTCTTAATAAATTTTGGC	657
Db	53883	CAATGCAAACTGGTTTATATTAGATACGTGACTTTAAATACTCAATACAGCTTTTCTCC	53824
Qy	658	A 658	
Db	53823	A 53823	

RESULT 9	HSJ193G15/c
LOCUS	HSJ193G15
DEFINITION	Human DNA sequence from clone 193G15 on chromosome 4. Contains the 5' part of the RGS12 gene for regulator of G-protein signalling 12, an RPL7A (60S Ribosomal Protein 7A (SURF3)) pseudogene, ESTs, an STS, GSSs and a putative CpG island, complete sequence.
ACCESSION	AL096838
VERSION	AL096838.1
KEYWORDS	HTG; 60S Ribosomal Protein 7A; CpG Island; G-protein; RGS12; RPL7A; SURF3.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 62493)
 AUTHORS McMurray, A.
 TITLE Direct Submission
 JOURNAL Submitted (10-JUL-1999) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 COMMENT requests: clonerequest@sanger.ac.uk
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em: EMBL; Sw: SWISSPROT; Tr: TREMBL
 IMPORTANT: This sequence is not the entire insert of clone 193G15.
 It may be shorter because we only sequence overlapping sections
 once, or longer because we arrange for a small overlap between
 neighbouring submissions.
 The true left end of clone 1E6 is at 6371 in this sequence. The
 true right end of clone 361H4 is at 104 in this sequence. This
 sequence has been finished according to sequence map criteria as
 follows. An attempt is made to resolve all sequencing problems,
 such as compressions and repeats, but not necessarily within known
 annotated human repeat sequence elements (e.g. Alu). Where the
 sequence is ambiguous, there is an annotation using the 'unsure'
 feature key.
 This sequence was generated from part of bacterial clone contigs of
 human chromosome 4, constructed by the Sanger Centre Chromosome 4
 Mapping Group.
 193G15 is from the library RPCI1 constructed at the Roswell Park
 Cancer Institute by the group of Pieter de Jong. For further
 details see <http://bacpac.med.buffalo.edu/VECTOR:pcYPAC2>.
 FEATURES
 Location/Qualifiers
 1..62493
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="4"
 /clone="RPI-193G15"
 /clone_lib="RPCI-1"
 17336 a 16390 c 13580 g 15187 t
 BASE COUNT
 ORIGIN

Query Match 36.6%; Score 245; DB 9; Length 62493;
 Best Local Similarity 70.9%; Pred. No. 6.9e-57;
 Matches 426; Conservative 0; Mismatches 110; Indels 65; Gaps 5;

QY 58 ATGTTCACTGGGCACTTCCCTTCACACCCCTTGGCCACGCTGGTACCGCTGGGGAGCTG 117
 Db 54358 AGGTTCACTGGGTGCTTCTCTCTGGCCCTTGGCCACGCTGGTAAACACTGTGGAGCTG 54299

QY 118 TGAGAGTGTGAGGGGACGTTCCAGCCGTCTGAGACTCTTTCTCTCTACTCAGACGCGAGC 177
 Db 54298 TGAGGGTGTGAGGGTGTGGTCTCTGCTTCTGAGCTCTTTACCCACTACTCAGAGCAGC 54239

QY 178 CTATAGTTCGGCAGGCGAGCTCTCCAGGAAGTGAATAGTGAATATGATTTGGCGAGG 237
 Db 54238 CGTT-----TGAAATATGATTTGGCAAGG 54215

QY 238 AAGATCAACATATAGGCTTAGCCCAAGAGAGTTTACAGCCTCTCTGAGCTGATTGGGGC 297
 Db 54214 AAGATCAGCATATAGGCTTAGCCCAAGAGAGCTGTATCTGAGCCTCTCTGAGCTAACTGGGCC 54155

QY 298 TATGCTTACTGCTCCCTTTGTCCAGGAACCCACTGATGAGAGCCCTAAAGAGAGAA 357
 Db 54154 TGTGCTT-----GAGCCAGTGTATGAGCAGCCTCAGCAGAGGA 54116

QY 358 ACCACCCACTAAAGTTCGGAATCTTACACCTGATCAGAGAGAGAGATGATCAGGGTGC 417
 Db 54115 ACCACCCACTGAAAGTTCGGGTCTT--ACCTGGCCAGGAGAGAGAGATCAGGTGC 54058

QY 418 ACCTAGATTAAGTGTGCTGCTGAGTGAAGCCGATCTCAGGAGCTATGTCAGACAAAGAC 477
 Db 54057 TGTGAGATTTCTGTGCTTGACAGGAAGCTGATCTCCGGAGCTGTCTCAAGA-----54003

QY 478 TGGGATGATGTGAAGTGTGCTGATGTCAAGGGAAGATTTTACCAAAGCAGAGCA 537
 Db 54002 CTGGGATGATGTGAGATGCTCTGATGTC--CGGGATATTTCTGCGAATCAGAGCA 53944

QY 538 CTTTAAATGCCAGAGCAGGTGAAGGGAATCAAGCTTTAAAGAAAGATAAGCTGAAA 597
 Db 53943 ATTAAATTTGCCAGAGGAGGTGAAGGGCAACACAGGTTCGAAGGAGACAAGCTGAAA 53884

QY 598 CAACCAAACTCTTTTATATATAGATATTTTACTTTAAAGAGTCTTAAATATTTTGGC 657
 Db 53883 CAATGCAAACTGTTTATATATAGATGCTGATTTAAATATCTCAATACAGTTTCTCC 53824

QY 658 A 658
 Db 53823 A 53823

RESULT 10
 AL645949/c
 LOCUS Human DNA sequence from clone RP11-357G3 on chromosome 4, complete
 DEFINITION sequence.
 ACCESSION AL645949
 VERSION AL645949.16 GI:24366471
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo Sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 145151)
 Brown.A.
 Direct Submission
 Submitted (16-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk
 On Oct 24, 2002 this sequence version replaced gi:24210295.
 COMMENT
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; SW; SWISPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP database can be found at
 http://www.sanger.ac.uk/Projects/_C_elegans/wormpep/ RP11-357G3 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBACe3.6.

FEATURES

Location/Qualifiers

source

1..145151
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosomes="4"
 /clone="RP11-357G3"
 /clone_lib="RP11-11.2"

BASE COUNT 39979 a 37603 c 33739 g 33830 t

ORIGIN

Query Match 36.6%; Score 245; DB 9; Length 145151;
 Best Local Similarity 70.9%; Pred. No. 6.9e-57;
 Matches 426; Conservative 0; Mismatches 110; Indels 65; Gaps 5;

QY 58 ATGTTCACTGGGCACTTCCCTTCACACCCCTTGGCCACGCTGGTACCGCTGGGGAGCTG 117
 Db 102391 AGGTTCACTGGGTGCTTCTCTCTGGCCCTTGGCCACGCTGGTAAACACTGTGACCTG 102332

QY 118 TGAGAGTGTGAGGGGACGTTCCAGCCGTCTGAGACTCTTTCTCTCTACTCAGACGCGAGC 177
 Db 102331 TGAGGGTGTGAGGGTGTGGTCTCTGCTTCTGAGCTCTTTACCCACTACTGAGAAAGCAGC 102272

QY 178 CTATAGTTCGGCAGGCGAGCTCTCCAGGAAGTGAATAGTGAATATGATTTGGCGAGG 237
 Db 102271 CGTT-----TGAAATATGATTTGGCAAGG 102248

QY 238 AAGATCAACATATAGGCTTAGCCCAAGAGAGTTTACAGCCTCTCTGAGCTGATTGGGGC 297
 Db 102247 AAGATCAGCATATAGGCTTAGCCCAAGAGAGCTGTATGTCAGCCTCTCTGAGCTAACTGGGGC 102188

QY 298 TATGCTTACTGCTCCCTTTGTCCAGGAACCCACTGATGAGAGCCCTAAAGAGAGAA 357
 Db 102187 TGTGCTT-----GAGCCAGTGTATGAGCAGCCTCAGCAGAGGA 102149

QY 358 ACCACCCACTAAAGTTCGGAATCTTACACCTGATCAGAGAGAGAGATGATCAGGGTGC 417
 Db 102148 ACCACCCACTGAAAGTTCGGGTCTT--ACCTGGCCAGGAGAGAGAGAGATCAGGTGC 102091

QY 418 AGCTGAGATCAAGTGCCTGAGCTGGAGCCATCTCCAGGAGCTATGTCAGACAAAGAC 477
 Db 102090 TGTGAGATTTCTGTCCTTGGCAGGAGCTGATCTCCGGGAGCTGTCTCAAGA-----102036

QY 478 TGGGATGATGTGAAGGTGGTACTGATGTCAGGGGAGAGATTCTACCAAAAGCAGAGCA 537
 Db 102035 CTGGATGATGTGGAGATGGTCTCTGATGTC--CGGGATATTTCTGCCGAATCAGAGCA 101977

QY 538 CTTTAAATGCCAGAGCAGGTGAAGGGAATCAAGCTTTAAAGAGAGATTAAGCTGAAA 597
 Db 101977 CTTTAAATGCCAGAGCAGGTGAAGGGAATCAAGCTTTAAAGAGAGATTAAGCTGAAA 597

Db 101976 ATTTAAATTGCGAGGAGGTGAAGCGCACACAGGTTCCAGAGGACAGCTGAAA 101917
 QY 598 CAACACAACTGTTTATATATAGATATTTTACCTTAAAGAGTCTTAAATATTTGGC 657
 Db 101916 CAATGCAAACTGGTTTATATATAGATAGCTTAAATATATCTCAATACAGTTTCTCC 101857
 QY 658 A 658
 Db 101856 A 101856

RESULT 11
 HSA318891 2648 bp DNA linear PRI 15-JAN-2002
 LOCUS HSA318891
 DEFINITION Homo sapiens partial XAGE-2 gene for XAGE-2 5' part, exons 1-2 (and joined CDS).
 ACCESSION AJ318891
 VERSION AJ318891.1 GI:18182816
 KEYWORDS XAGE-2 gene.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Zandman, A.J.W., van Kraats, A.A., Weidle, U.H., Ruiter, D.J. and Van
 Muijen, G.N.P.
 Expression profile and alignment of the XAGE family of
 cancer/testis associated genes
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2648)
 AUTHORS Zandman, A.J.W.
 TITLE Direct Submission
 JOURNAL Submitted (15-AUG-2001) Zandman A.J.W., Department of Pathology,
 University Medical Center, Geert Grooteplein Zuid 24, P.O. Box 9101
 6500 HB Nijmegen, NETHERLANDS
 COMMENT Related sequence AJ318892.
 FEATURES
 1. 2648 Location/Qualifiers
 source
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="X"
 /map="Xp11.21-Xp11.22"
 join(1441..2601,AJ318892.1:407..5517)
 /gene="XAGE-2"
 join(1441..1720,2513..2601,AJ318892.1:407..512,
 AJ318892.1:2384..2509,AJ318892.1:5409..5517)
 /label="XAGE-2"
 1441..1720
 /gene="XAGE-2"
 /number=1
 1721..2512
 /gene="XAGE-2"
 /number=1
 2513..2601
 /gene="XAGE-2"
 /number=2
 join(2521..2601,AJ318892.1:407..512,AJ318892.1:2384..2509,
 AJ318892.1:5409..5431)
 /gene="XAGE-2"
 /codon_start=1
 /label="XAGE-2_CDS"
 /product="XAGE-2 protein"
 /protein_id="CAC88125.1"
 /db_xref="GI:18182817"
 /translation="NSWRGRSTYRPRRSLQPELIGAMLEPTDEPKEKPTKSR
 NPTDKRRDQGAETIQVPLEADQLCQTKTGDCGCGTGVKGLPKAEHFKNP
 EAGGKQSV"

BASE COUNT 731 a 557 c 717 g 643 t
 ORIGIN
 Query Match 32.3%; Score 216; DB 9; Length 2648;

Best Local Similarity 93.8%; Pred. No. 6.1e-49;
 Matches 225; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 1 CACAACGCGAGGACCGACTTCAGTGTGCAATTCCTTGGACACCTGCTCAGTGTGCATG 60
 Db 1506 CACAACGCGAGGACCGACTTCAGTGTGCAATTCCTTGGACACCTGCTCAGTGTGCATG 1565
 QY 61 TTCACCTGGGCATCTTCCCTTTCGACCCCTTGGCCACGTGGTGACCGCTGGGGAGCTGTGA 120
 Db 1566 TTCACCTGGGCATCTTCCCTTTCGACCCCTTGGCCACGTGGTGACCGCTGGGGAGCTGTGA 1625
 QY 121 GAGTGTGAGGCGGACGCTTCCAGCCGCTCTGGACTCTTTCTCTCTACTGAGACGCGCCTTA 180
 Db 1626 GAGTGTGAGGCGGACGCTTCCAGCCGCTCTGGACTCTTTCTCTACTGAGACGCGCCTTA 1685
 QY 181 TAGGTCCGCGAGCCAGTCTCTCCAGCAACTGAAATAGTGAATATAGTTCGCGAGGAAG 240
 Db 1686 TAGGTCCGCGAGCCAGTCTCTCCAGCAACTGAAATAGTGAATATAGTTCGCGAGGAAG 1745

RESULT 12
 AL953874
 LOCUS AL953874
 DEFINITION Human DNA sequence from clone CTD-2267G17 on chromosome X, complete
 sequence.
 ACCESSION AL953874
 VERSION AL953874.2 GI:24527895
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 94514)
 Direct Submission
 Submitted (01-NOV-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk
 On Nov 3, 2002 this sequence version replaced gi:24410496.
 COMMENT

Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at
 http://www.sanger.ac.uk/projects/C.elegans/wormpep
 This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/ChrX
 CTD-2267G17 is from the Caltech genomic sperm BAC library D.
 VECTOR: pBelosAC11.

FEATURES
 source
 1..94514
 /organism="Homo sapiens"
 /mol_type="genomic DNA"

chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChrX>. RP11-472D17 is from the library RPC1-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>. VECTOR: pBACe3.6.

```

FEATURES
  source
    1. 121324
      Location/Qualifiers
        organism="Homo sapiens"
        mol_type="genomic DNA"
        db_xref="taxon:9606"
        chromosome="X"
        clone="RP11-472D17"
        clone_lib="RP11-11.2"
    36831 a 25068 c 24273 g 35152 t
BASE COUNT

```

```

BASE COUNT
36831 a 25068 c 24273 q 35152 t
/db xref="taxon:9606"
/chromosome="X"
/clone="RP11-472D17"
/clone_lib="RP11-11.2"

```

```

BASE COUNT      36831 a 25068 c 24273 g 35152 t
ORIGIN
Query Match      32.3%      Score 216;  DB 9;  Length 121324;
Best Local Similarity 93.8;  Pred. No. 8.2e-49;
Matches 225;  Conservative 0;  Mismatches 15;  Indels 0;  Gaps 0;

```

Query Match 52.35; Score 210, 8.2e-49; Mongol 225; 0; Gaps 0;
Best Local Similarity 93.8%; Pred. No. 8.2e-49;
Matches 225; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
1 CACACCGAGCCGACCTTACGTGCGCATGTCTCTTGACACCTGCGCTCAGTGGCATG 60

Accession	Sequence	Length
QY	1 CACAAACGAGGACACCGACTTCAGTGTGCGATGTTCTTGGACACCTTGCCCTCAGTGTGCGATG	80
Db	11241 CACAAACGAGGACACCGACTTCAGTGTGCGATGTTCTTGGACACCTTGCCCTCAGTGTGCGATG	112
...	51 TTTCCATCTGACGCGATGTTCTTGGACACCTTGCCCTCAGTGTGCGATGTTCTTGGACACCTTGCCCTCAGTGTGCGATG	120

Oy

61 TTACTGGGCATCTTCCTTTCGACCCCTTTGCCCGCAGTGGTGACGCTGGGAGCTGTGA 120
| | | | |
Db

112381 TTCAC TGGGCATCTTCCTTTCGACCCCTTTGCCCGCAGTGGTGACGCTGGGAGCTGTGA 112
| | | | |

[illegible]

181 TAGGTCCCGAGGCGAGTCTCTCCGAGGAAC TGAATAATAGTGAATATAGTTCGCGAGGAG 240
112261 TAGGTCCCGAGGCGCACTCTCTCCGAGGAAC TGAATGTGAGTGAATGTGAGGAGGAGTGTGAG 112261

RESULT 14			
AX198429	AX198429	320 bp	DNA
LOCUS			linear
			PAT 29-AUG-2001

DEFINITION Sequence 884 from Patent WO0151513.
 ACCESSION AX138429
 VERSION AX138429.1 GI:15388750
 KEYWORDS .
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 Alga, P.A.
AUTHORS Ovarian tumor-associated sequences

```

FEATURES
    source
        1. 320
            Location/Qualifiers
                Patent: WO 0151513-A 894 19-JUL-2001;
                CORNIX CORPORATION (US)
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"

```

BASE COUNT	96 a	65 c	85 g	74 t
ORIGIN				
Query Match	32.1%	Score 215;	DB 6;	Length 320;
Best Local Similarity	100.0%;	Pred. No. 9.9e-49;		
Matches 215;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
Ov	420	CTGAGATTCAAGTGCCTGACCTGGAGCCGATCTCCAGAGCTATGTCAGACCAAGACTG	479	

1 CTGGATTCAAGTGCCTGACCTGGGAAGCGGATCTCCAGGAGCTATGTGACACAAGACTG 60
DB
480 GGGATGGATGTGAAGGTGCTACTGATCTCAAGGGGAGAGTTCACCAAAAGCAGAGCACT 5399
QY

Db 61 GGGATGGATGTAAGGTGGTACTGATGTCAGGGGAAGATTCTACCAAAAGCAGAGCACT 120
Qy 540 TTAATAATGCCAGACAGGTGARGGGAAATCACAGGTTTAAAGGAAGATAAGCTGAAACA 599
Db 121 TTAATAATGCCAGACAGGTGARGGGAAATCACAGGTTTAAAGGAAGATAAGCTGAAACA 180
Qy 600 ACACAAACTGTTTATATATAGATATTTACTTTA 634
Db 181 ACACAAACTGTTTATATATAGATATTTACTTTA 215
RESULT 15
LOCUS AX208978 214 bp DNA linear PAT 31-AUG-2001
DEFINITION Sequence 818 from Patent WO0157207.
ACCESSION AX208978
VERSION AX208978.1 GI:15423401
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Algate, P.A. and Mannion, J.
TITLE Compositions and methods for the therapy and diagnosis of ovarian cancer
JOURNAL Patent: WO 0157207-A 818 09-AUG-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..214
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 74 a 34 c 55 g 51 t
ORIGIN
Query Match 32.0%; Score 214; DB 6; Length 214;
Best Local Similarity 100.0%; Pred. No. 1.8e-48;
Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 420 CTGAGATTCAAGTGGCTGACCTGGAGCGGATCTCAGGAGCTATGTGAGCAAGACTG 479
Db 1 CTGAGATTCAAGTGGCTGACCTGGAGCGGATCTCAGGAGCTATGTGAGCAAGACTG 60
Qy 480 GGGATGGATGTAAGGTGGTACTGATGTCAGGGGAAGATTCTACCAAAAGCAGAGCACT 539
Db 61 GGGATGGATGTAAGGTGGTACTGATGTCAGGGGAAGATTCTACCAAAAGCAGAGCACT 120
Qy 540 TTAATAATGCCAGACAGGTGARGGGAAATCACAGGTTTAAAGGAAGATAAGCTGAAACA 599
Db 121 TTAATAATGCCAGACAGGTGARGGGAAATCACAGGTTTAAAGGAAGATAAGCTGAAACA 180
Qy 600 ACACAAACTGTTTATATATAGATATTTACTTTT 633
Db 181 ACACAAACTGTTTATATATAGATATTTACTTTT 214
RESULT 16
LOCUS AR272357 399 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 69 from patent US 6504010.
ACCESSION AR272357
VERSION AR272357.1 GI:29704242
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 399)
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S., Carter, D., Retter, M.W., Mannion, J., and Fan, L.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: US 6504010-A 69 07-JAN-2003;
FEATURES Location/Qualifiers
source 1..399
/organism="unknown"
BASE COUNT 150 a 87 c 94 g 67 t 1 others
ORIGIN
Query Match 30.5%; Score 204; DB 6; Length 399;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
Qy 319 GTCCAGGAAACCCACTGATGAAGAGCCCTAAAGAGAGAAACCCACCCTAAAGTCGGAA 378
Db 40 GTCCAGGAGCCCGAGTAAATGGAGAGCCCAAAAGAGAGAACCCAGCAGCTGAAAGTCGGGA 99
Qy 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGGAGCTGAGATTCAAGTGCCTGA 438
Db 100 TCCTACACCTGGGAGAGAGAGAGATGATCAGGGTGGAGCTGAGATTCAAGTGCCTGA 159
Qy 439 CTTGGAAGCGGATCTCCAGGAGCTATGTGAGCAACAAGACTGGGGATGGATGGAAGGT-G 497
Db 160 CATGGAGAGTGTCTGCAAGAGCTGCATCAGTCAAAACCCGGGATAAATCTGGATTGG 219
Qy 498 GTRACTGATGTCAAGGGGAAAGATTCTACCAAAAGCAGAGCACTTAAATGCCAGAGCAG 557
Db 220 GTTCCGGCTCAAGGTGAAGATAATACCTTAAAGAGAGAACACTGTAAATGCCAGAGCAG 279
Qy 558 GTCAAGGGAATCACAGGTTTAAAGGAAGATGAAGCTGAAACCAACAACTGTTTTATA 617
Db 280 GTGAGAGCAACCAAGTTTAAATGAAGCAGCTGAAACCAACCGCAAGCTGTTTTATA 339
Qy 618 TTAGATATTTACTTTTAAAGAGCTCTTAAATTTT 653
Db 340 TTAGATATTTACTTTTAAAGAGCTCTTAAATTTT 375
RESULT 17
LOCUS AR275938 399 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 69 from patent US 6509448.
ACCESSION AR275938
VERSION AR275938.1 GI:29709583
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 399)
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S., Carter, D., Retter, M.W., Mannion, J., Fan, L., and Wang, A.
TITLE Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL Patent: US 6509448-A 69 21-JAN-2003;
FEATURES Location/Qualifiers
source 1..399
/organism="unknown"
BASE COUNT 150 a 87 c 94 g 67 t 1 others
ORIGIN
Query Match 30.5%; Score 204; DB 6; Length 399;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
Qy 319 GTCCAGGAAACCCACTGATGAAGAGCCCTAAAGAGAGAAACCCACCCTAAAGTCGGAA 378
Db 40 GTCCAGGAGCCCGAGTAAATGGAGAGCCCAAAAGAGAGAACCCAGCAGCTGAAAGTCGGGA 99
Qy 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGGAGCTGAGATTCAAGTGCCTGA 438
Db 100 TCCTACACCTGGGAGAGAGAGAGATGATCAGGGTGGAGCTGAGATTCAAGTGCCTGA 159
Qy 439 CTTGGAAGCGGATCTCCAGGAGCTATGTGAGCAACAAGACTGGGGATGGATGGAAGGT-G 497
Db 160 CATGGAGAGTGTCTGCAAGAGCTGCATCAGTCAAAACCCGGGATAAATCTGGATTGG 219

```
QY 498 GTACTGATGTCAGGGGAGATTCTACCAAGACGAGACACTTTAAATGCCAGAGCAG 557
Db 220 GTTCCGGGCTCAAGCTGAAGTAATACCTTAAGAGGAACTGTAAATGCCAGAGCAG 279
QY 558 GTGAAGGGAATCAGAGGTTTAAAGGAGATAGCTGAAACAAACAACTGTTTTTATA 617
Db 280 GTGAAGAGCAACCAAGTTTAAATGAAGACAGCTGAAACCAACGACAGCTGGTTTTATA 339
QY 618 TTAGATATTTTACTTTAAAGAGCTTTAATAAATTTT 653
Db 340 TTAGATATTTGACTTAAACTATCTCAATAAAGTTTT 375

RESULT 18
AX062442 399 bp DNA linear PAT 24-JAN-2001
LOCUS
DEFINITION Sequence 69 from Patent WO0100828.
ACCESSION AX062442
VERSION AX062442.1 GI:12540317
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1.
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,
Carter, D., Retter, M.W. and Mannion, J.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0100828-A 69 04-JAN-2001;
CORIXA CORPORATION (US)
FEATURES
Location/Qualifiers
source
1..399
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 150 a 87 c 94 g 67 t 1 others
ORIGIN
Query Match 30.5%; Score 204; DB 6; Length 399;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGAAACCCACTGATGAAGAGCTTAAGAGAGAACCCACCCACTAAAGTCGGA 378
Db 40 GTCCAGGAGGCCAGTAAATGGAGAGCCCAAAAGAGAACCCAGCAGCTGAAGTCGGA 99
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGA 438
Db 100 TCCTACACCTGGCAGCAGACAGAGAGATCAGATACAGCTGAGATCCAGTGCAGCA 159
QY 439 CTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGCTGGGGATGGATGGAAGT-G 497
Db 160 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGGATAAATCTGATTTGG 219
QY 498 GTACTGATGTCAGGGGAGATTCTACCAAGACGAGACACTTTAAATGCCAGAGCAG 557
Db 220 GTTCCGGGCTCAAGCTGAAGTAATACCTTAAGAGGAACTGTAAATGCCAGAGCAG 279
QY 558 GTGAAGGGAATCAGAGCTTTAAAGGAGATAGCTGAAACAAACAACTGTTTTTATA 617
Db 280 GTGAAGAGCAACCAAGTTTAAATGAAGACAGCTGAAACCAACGACAGCTGGTTTTATA 339
QY 618 TTAGATATTTTACTTTAAAGAGCTTTAATAAATTTT 653
Db 340 TTAGATATTTGACTTAAACTATCTCAATAAAGTTTT 375

RESULT 19
AX367359 399 bp DNA linear PAT 16-FEB-2002
LOCUS
DEFINITION Sequence 69 from Patent WO0204514.
ACCESSION AX367359
```

```
VERSION AX367359.1 GI:18855463
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1.
AUTHORS Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W.,
Mannerakis, M., Carter, D., Fanger, G.R., Vedvick, T.S., Bangur, C.S.,
McNabb, A., Fanger, N., Switzer, A., McNeill, P.D. and Clapper, J.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0204514-A 69 17-JAN-2002;
CORIXA CORPORATION (US)
FEATURES
Location/Qualifiers
source
1..457
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 150 a 87 c 94 g 67 t 1 others
ORIGIN
Query Match 30.5%; Score 204; DB 6; Length 399;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGAAACCCACTGATGAAGAGCTTAAGAGAGAACCCACCCACTAAAGTCGGA 378
Db 40 GTCCAGGAGGCCAGTAAATGGAGAGCCCAAAAGAGAACCCAGCAGCTGAAGTCGGA 99
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGA 438
Db 100 TCCTACACCTGGCAGCAGACAGAGAGATCAGATACAGCTGAGATCCAGTGCAGCA 159
QY 439 CTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGCTGGGGATGGATGGAAGT-G 497
Db 160 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGGATAAATCTGATTTGG 219
QY 498 GTACTGATGTCAGGGGAGATTCTACCAAGACGAGACACTTTAAATGCCAGAGCAG 557
Db 220 GTTCCGGGCTCAAGCTGAAGTAATACCTTAAGAGGAACTGTAAATGCCAGAGCAG 279
QY 558 GTGAAGGGAATCAGAGCTTTAAAGGAGATAGCTGAAACAAACAACTGTTTTTATA 617
Db 280 GTGAAGAGCAACCAAGTTTAAATGAAGACAGCTGAAACCAACGACAGCTGGTTTTATA 339
QY 618 TTAGATATTTTACTTTAAAGAGCTTTAATAAATTTT 653
Db 340 TTAGATATTTGACTTAAACTATCTCAATAAAGTTTT 375

RESULT 20
AX273057 457 bp DNA linear PAT 10-APR-2003
LOCUS
DEFINITION Sequence 790 from patent US 6504010.
ACCESSION AR273057
VERSION AR273057.1 GI:29704942
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 457)
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,
Carter, D., Retter, M.W., Mannion, J. and Fan, L.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: US 6504010-A 790 07-JAN-2003;
FEATURES
Location/Qualifiers
source
1..457
/organism="unknown"
BASE COUNT 146 a 105 c 121 g 85 t
ORIGIN
```

Query Match 30.5%; Score 204; DB 6; Length 457;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

QY 319 GTCCAGGAAACCACTGATGAAGAGCCCTAAAGAGAGAGAAACACCCACTAAAGAGTCGGAA 378
DB 120 GTCCAGGAGCCAGTAATGGAGAGCCCAAAAGAGAACACGACGCTGAAGTCGGGA 179

QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCAGGTGAGATTCAGTGCCTGA 438
DB 180 TCCTACACCTGGGACGACAGAGAGAGATCAGGATCAGGTGAGATTCAGTGCCTGA 239

QY 439 CTTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGGATGATGTGAAGGT-G 497
DB 240 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAGTCAACACCGGGGATAATCTGGATTGG 299

QY 498 GTACTGATGTCAGGAGAGATTTACCAAAAGAGAGACATTTTAAATGCCAGAGCAG 557
DB 300 GTTCCGGCTCAAGGTGAAGATAATACCTAAAGAGAGAACTGTGTAATGCCAGAGCAG 359

QY 558 GTGAGGGAATACACAGGTTTAAAGAGAGATGAAGTGAACAAACAACTGTTTTTATA 617
DB 360 GTGAGAGACACACAGTTTAAATGAAGACAGCTGAACACCAAGCTGGTTTTTATA 419

QY 618 TTAGATATTTTACTTTAAAGAGCTTTAATAAATTTT 653
DB 420 TTAGATATTTGACTTAAACTATCTCAATAAAGTTTT 455

RESULT 21
LOCUS AR276638 457 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 790 from patent US 6509448.
ACCESSION AR276638
VERSION AR276638.1 GI:29710285
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 457)
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedwick, T.S.,
Carter, D., Retter, M.W., Mannion, J., Fan, L., and Wang, A.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: US 6509448-A 790 21-JAN-2003;
FEATURES
source Location/Qualifiers
1. .457
/organism="unknown"
BASE COUNT 146 a 105 c 121 g 85 t
ORIGIN

Query Match 30.5%; Score 204; DB 6; Length 457;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

QY 319 GTCCAGGAAACCACTGATGAAGAGCCCTAAAGAGAGAGAAACACCCACTAAAGAGTCGGAA 378
DB 120 GTCCAGGAGCCAGTAATGGAGAGCCCAAAAGAGAACACGACGCTGAAGTCGGGA 179

QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCAGGTGAGATTCAGTGCCTGA 438
DB 180 TCCTACACCTGGGACGACAGAGAGAGATCAGGATCAGGTGAGATTCAGTGCCTGA 239

QY 439 CTTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGGATGATGTGAAGGT-G 497
DB 240 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAGTCAACACCGGGGATAATCTGGATTGG 299

QY 498 GTACTGATGTCAGGAGAGATTTACCAAAAGAGAGACATTTTAAATGCCAGAGCAG 557
DB 300 GTTCCGGCTCAAGGTGAAGATAATACCTAAAGAGAGAACTGTGTAATGCCAGAGCAG 359

QY 558 GTGAGGGAATACACAGGTTTAAAGAGAGATGAAGTGAACAAACAACTGTTTTTATA 617
DB 360 GTGAGAGACACACAGTTTAAATGAAGACAGCTGAACACCAAGCTGGTTTTTATA 419

QY 618 TTAGATATTTTACTTTAAAGAGCTTTAATAAATTTT 653
DB 420 TTAGATATTTGACTTAAACTATCTCAATAAAGTTTT 455

Db 360 GTGAAGAGCAACCAACCAAGTTTAAATGAAGACAAAGCTGAACCAACGCAAGCTGGTTTTATA 419

QY 618 TTAGATATTTTACTTTAAAGAGCTTTAATAAATTTT 653

Db 420 TTAGATATTTGACTTAAACTATCTCAATAAAGTTTT 455

RESULT 22
LOCUS AX063163 457 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 790 from Patent WO0100828.
ACCESSION AX063163
VERSION AX063163.1 GI:12541017
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedwick, T.S.,
Carter, D., Retter, M.W., and Mannion, J.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0100828-A 790 04-JAN-2001;
FEATURES
source Location/Qualifiers
1. .457
/organism="Homo sapiens"
/mol type="genomic DNA"
/db xref="taxon:9606"
BASE COUNT 146 a 105 c 121 g 85 t
ORIGIN

Query Match 30.5%; Score 204; DB 6; Length 457;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

QY 319 GTCCAGGAAACCACTGATGAAGAGCCCTAAAGAGAGAGAAACACCCACTAAAGAGTCGGAA 378
DB 120 GTCCAGGAGCCAGTAATGGAGAGCCCAAAAGAGAACACGACGCTGAAGTCGGGA 179

QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCAGGTGAGATTCAGTGCCTGA 438
DB 180 TCCTACACCTGGGACGACAGAGAGAGATCAGGATCAGGTGAGATTCAGTGCCTGA 239

QY 439 CTTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGGATGATGTGAAGGT-G 497
DB 240 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGGATAATCTGGATTGG 299

QY 498 GTACTGATGTCAGGAGAGATTTACCAAAAGAGAGACATTTTAAATGCCAGAGCAG 557
DB 300 GTTCCGGCTCAAGGTGAAGATAATACCTAAAGAGAGAACTGTGTAATGCCAGAGCAG 359

QY 558 GTGAAGGGAATACACAGGTTTAAAGAGAGATGAAGTGAACAAACAACTGTTTTTATA 617
DB 360 GTGAAGAGCAACCAACCAAGTTTAAATGAAGACAAAGCTGAACCAACGCAAGCTGGTTTTATA 419

QY 618 TTAGATATTTTACTTTAAAGAGCTTTAATAAATTTT 653
DB 420 TTAGATATTTGACTTAAACTATCTCAATAAAGTTTT 455

RESULT 23
LOCUS AX368080 457 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 790 from Patent WO0204514.
ACCESSION AX368080
VERSION AX368080.1 GI:18856163
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE	1	Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W., Marnerakis, M., Carter, D., Fanger, G.R., Vedvick, T.S., Bangur, C.S., McNabb, A., Fanger, N., Switzer, A., McNeill, P.D. and Clapper, J.D.
AUTHORS		Compositions and methods for the therapy and diagnosis of lung cancer
TITLE		Patent: WO 0204514-A 790 17-JAN-2002;
JOURNAL		CORIXA CORPORATION (US)
FEATURES		Location/Qualifiers
source		1. .457
BASE COUNT	146 a	105 c 121 g 85 t
ORIGIN		
Query Match	30.5%;	Score 204; DB 6; Length 457;
Best Local Similarity	77.4%;	Pred. No. 1.2e-45;
Matches	260; Conservative	0; Mismatches 75; Indels 1; Gaps 1;
QY	319	GTCCAGGACCCACTGATGAAGAGCCTAAAGAGAGAACCCACCCTAAAGTCGGAA 378
Db	120	GTCCAGGACCCAGTAATGGAGGCCCCCAAAAGAGAACCCAGCAGCTGAAGTCGGGA 179
QY	379	TCTACACCTGATCAGAGAGAGAGATGATCAGGGTGAGCTGAGATTCAGAGTCGCTGA 438
Db	180	TCTACACCTGGGCAGCAGACAGAGATCAGGATACAGCTGAGATCCAGTCGGCGA 239
QY	439	CTGGAGCGGCTCTCCAGAGCTATGTCAGACAAAGCTGGGGATGATGAGAGGT-G 497
Db	240	CATGGAAGTGATCTCCAGAGCTGATCAGTCAACACCCGGGATAAATCTGGATTGG 299
QY	498	GTACTGATGTCAGGGGAAGATTCTACCAAAAGCAGAGCACTTTAAATGCCAGAAGCAG 557
Db	300	GTTCGGCGTCAAGGTGAAGATAATACCTAAAGAGGAACACTGTAAATGCCAGAAGCAG 359
QY	558	GTGAAGGGAATCAGAGGTTTAAAGGAGTAAGCTGAACACACAACTCTTTTATA 617
Db	360	GTGAAGGAGCAACCAAGTTTAAATGAAGCAAGCTGAACACCAAGCTGTTTATA 419
QY	618	TTAGATATTCTTAAAGAGTCTTAATAAATTTT 653
Db	420	TTAGATATTGACTTAAACTATCTCAATAAGTTTT 455
RESULT 24		
AX370580		
LOCUS	AX370580	463 bp DNA linear PAT 01-MAR-2002
DEFINITION	Sequence 1 from Patent EP1179589.	
ACCESSION	AX370580	
VERSION	AX370580.1	GI:19168789
KEYWORDS		Homo sapiens (human)
SOURCE		Homo sapiens
ORGANISM		Homo sapiens
REFERENCE		
AUTHORS		
TITLE		Mmx-1, a member of the family of human cancer/testis antigens, a protein encoded thereby and a process for determining whether a tumor sample has metastatic potential
JOURNAL		Patent: EP 1179589-A 1 13-FEB-2002;
FEATURES		F. HOFFMANN-LA ROCHE AG (CH)
source		Location/Qualifiers
1. .463		
/organism="Homo sapiens"		
/mol_type="genomic DNA"		
/db_xref="taxon:9606"		
126. .371		
/note="unnamed protein product"		
/codon_start=1		
/protein_id="CAD26691.1"		
/db_xref="GI:19168790"		
CDS		
translation="MESPKKNQQLKVGILHLGSRQKKIRIQIRSQCATWKVICSCI		
SOTFGINLDLGSQGVKKIIPKEEHCKMPAGESEQPV"		
154 a	107 c	120 g 82 t
BASE COUNT		
ORIGIN		
Query Match	30.5%;	Score 204; DB 6; Length 463;
Best Local Similarity	77.4%;	Pred. No. 1.2e-45;
Matches	260; Conservative	0; Mismatches 75; Indels 1; Gaps 1;
QY	319	GTCCAGGACCCACTGATGAAGAGCCTAAAGAGAGAACCCACCCTAAAGTCGGAA 378
Db	109	GTCCAGGAGCCAGTAATGGAGGCCCCCAAAAGAGAACCCAGCAGCTGAAGTCGGGA 168
QY	379	TCTACACCTGATCAGAGAGAGAGATGATCAGGGTGAGCTGAGATTCAGAGTCGCTGA 438
Db	169	TCTACACCTGGCAGCAGACAGAGAGATCAGGATACAGCTGAGATCCAGTCGGCGA 228
QY	439	CTGGAGCGGATCTCCAGAGCTATGTCAGACAAAGCTGGGGATGATGAGAGGT-G 497
Db	229	CATGGAAGTGATCTCCAGAGCTGATCAGTCAACACCCGGGATAAATCTGGATTGG 288
QY	498	GTACTGATGTCAGGGGAAGATTCTACCAAAAGCAGAGCACTTTAAATGCCAGAAGCAG 557
Db	289	GTTCGGCGTCAAGGTGAAGATAATACCTAAAGAGGAACACTGTAAATGCCAGAAGCAG 348
QY	558	GTGAAGGGAATCAGAGGTTTAAAGGAGTAAGCTGAACACACAACTGTTTATA 617
Db	349	GTGAAGGAGCAACCAAGTTTAAATGAAGCAAGCTGAACACCAAGCTGTTTATA 408
QY	618	TTAGATATTCTTAAAGAGTCTTAATAAATTTT 653
Db	409	TTAGATATTGACTTAAACTATCTCAATAAAGTTTT 444
RESULT 25		
HSA290447		
LOCUS	HSA290447	463 bp mRNA linear PRI 05-MAY-2001
DEFINITION	Homo sapiens mRNA for 9 kD cancer/testis associated protein (Xage-1b).	
ACCESSION	AJ290447.1	GI:13992498
VERSION	AJ290447	
KEYWORDS		cancer/testis associated protein; Ctp9 gene.
SOURCE		Homo sapiens (human)
ORGANISM		Homo sapiens
REFERENCE		
AUTHORS		Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE		Zendman, A.J.W., van Kraats, A.A., Weidle, U.H., Ruiter, D.J. and Van Muijen, G.N.P.
JOURNAL		CTP9, a novel cancer/testis-associated gene, induced in melanoma metastases
REFERENCE		2 (bases 1 to 463)
AUTHORS		Unpublished
TITLE		Direct Submission
JOURNAL		Submitted (04-MAY-2000) Zendman A.J.W., Pathology, University Medical Center St. Radboud, Geert Grooteplein Zuid 24, Nijmegen, 6500 HB POBOX 9101, NETHERLANDS
REMARK		Revised by author 22-JAN-2001
FEATURES		Location/Qualifiers
source		1. .463
/organism="Homo sapiens"		
/mol_type="mRNA"		
/db_xref="taxon:9606"		
/chromosome="X"		
/map="Xp11.21-Xp11.22"		
/tissue_type="melanoma metastasis"		
1. .463		
/gene="Xage-1b"		
125. .370		
/gene="Xage-1b"		
/codon_start=1		
/product="9 kD cancer/testis associated protein"		


```
/protein_id="CAC38107.1"
/db_xref="GI:1392499"
/translation="MESPKKKNQOLKVGILHLSRQKIRIQLRSQCATWKVICKSCI
SQTGINDLGSVGVKVIIPKECHKMPEAGEBPQV"
BASE COUNT      155 a   106 c   120 g      82 t
ORIGIN
Query Match      30.5%; Score 204; DB 9; Length 463;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGAACCCACTGATGAAGAGCCTAAAGAGAGAGAAACCCACCTAAAGTCGGA 378
Db 108 GTCCAGGAGCCAGTAATGAGAGCCCAAAAAGAGAACCCAGCAGCTGAAGTCGGA 167
QY 379 TCCTACACTGATCAGAGAGAGAGATGATCAGGTGCAGCTGAGATTCAAGTGCCTGA 438
Db 168 TCCTACACTGGGAGAGAGAGAGATGATCAGGTGCAGCTGAGATTCAAGTGCCTGA 227
QY 439 COTGAAGCCGATCTCCAGGAGCTATGTACAGAAAGAGCTGGGATGGATGAGGT-G 497
Db 228 CATGAAGGTGATCTGAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGG 287
QY 498 CTACTGATGTCNAGGGAGATTCTACCAAGAGAGAGCACTTTAAATGSCAGAGAG 557
Db 288 GTTCGGCGTCAAGGTGAAGATAATACCTAAAGAGAGAACTGTAAATGCCAGAGAG 347
QY 558 GTGAAGGGAATCAGAGGTTTAAAGAGAGATAAGCTGAACCAACACAACTGTTTTATA 617
Db 348 GTGAAGAGACCAACCAAGTTTAAATGAAGAGCAAGCTGAACCAACCGAGTGTGTTATA 407
QY 618 TTAGATATTTTACTTTAAAGAGCTCTTAATAATTTT 553
Db 408 TTAGATATTTGACTTTAAACTATCTCAATAAAGTTT 443

RESULT 26
BC009538      529 bp      mRNA      linear      PRI 22-OCT-2001
LOCUS      Homo sapiens, clone MGC:10063 IMAGE:3893227, mRNA, complete cds.
DEFINITION      BC009538
ACCESSION      BC009538
VERSION      BC009538.1 GI:16306934
KEYWORDS      MGC.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Strausberg,R.
Direct Submission
Submitted (29-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTB/DRP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 14 Row: 9 Column: 8.
Location/Qualifiers
1. .529
/organism="Homo sapiens"

FEATURES
source
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:10063 IMAGE:3893227"
/tissue_type="lung, carcinoma, large cell
undifferentiated."
/clone_lib="NIH MGC_69"
/lab_host="DH103"
/notes="Vector: pCMV-SPORT6"
184. .429
/codon_start=1
/product="Unknown (protein for MGC:10063)"
/protein_id="AAH09538.1"
/db_xref="GI:16306935"
/translation="MESPKKKNQOLKVGILHLSRQKIRIQLRSQCATWKVICKSCI
SQTGINDLGSVGVKVIIPKECHKMPEAGEBPQV"
BASE COUNT      184 a   116 c   146 g      83 t
ORIGIN
Query Match      30.5%; Score 204; DB 9; Length 529;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGAACCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCTAAAGTCGGA 378
Db 167 GTCCAGGAGCCAGTAATGAGAGCCCAAAAAGAGAACCCAGCAGCTGAAGTCGGA 226
QY 379 TCCTACACTGATCAGAGAGAGAGATGATCAGGTGCAGCTGAGATTCAAGTGCCTGA 438
Db 227 TCCTACACTGGGAGAGAGAGATGATCAGGTGCAGCTGAGATTCAAGTGCCTGA 286
QY 439 CTGGAGCGGATCTCCAGGAGCTATGTACAGAAAGAGCTGGGATGGATGGAAGGT-G 497
Db 287 CATGAAGGTGATCTGAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGG 346
QY 498 GTACTGATGTCAAGGGGAGATTCTACAAAGAGAGAGCACTTTAAATGCCAGAGAG 557
Db 347 GTTCGGCGTCAAGGTGAAGATAATACCTAAAGAGAGAACTGTAAATGCCAGAGAG 406
QY 558 GTGAAGGGAATCAGAGGTTTAAAGAGAGATGATGAAACCAACCACTGTTTTATA 617
Db 407 GTGAAGGCAACCAAGTTTAAATGAAGAGAGCTGAACCAACCGAGTGTGTTATA 466
QY 618 TTAGATATTTTACTTTAAAGAGCTCTTAATAATTTT 653
Db 467 TTAGATATTTGACTTTAAACTATCTCAATAAAGTTT 502

RESULT 27
AF251237      626 bp      mRNA      linear      PRI 23-AUG-2000
LOCUS      Homo sapiens XAGE-1 mRNA, complete cds.
DEFINITION      AF251237
ACCESSION      AF251237
VERSION      AF251237.1 GI:9885324
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 626)
XAGE-1, A New Gene That is Frequently Expressed in Ewing's Sarcoma
AUTHORS      Liu,X.F., Helman,L.J., Yeung,C., Lee,B. and Pastan,I.
TITLE      XAGE-1, A New Gene That is Frequently Expressed in Ewing's Sarcoma
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 626)
AUTHORS      Liu,X.F., Helman,L.J., Yeung,C., Lee,B. and Pastan,I.
TITLE      Direct Submission
JOURNAL      Submitted (31-MAR-2000) Lab of Mol. Biol., NCI, NIH, 37 Convent Dr.
MSC 4255, Bethesda, MD 20892, USA
Location/Qualifiers
1. .626
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X"
```

```
CDS
/tissue_type="testis"
86..526
/note="expressed in Ewing's sarcoma"
/codon_start=1
/product="XAGE-1"
/protein_id="XAGE-1"
/db_xref="GI:9885325"
/db_xref="GI:9885325"
/translation="MLLNCPPQCACSLGVFPSPVWGTTRSCPEPATRVPEVWLSP
LLRHGHGTQNTHTASPRSPVMSPKKNQQLKVLHLSGROKIRIOLRSQCATWK
VICKSCISQTPGINLDLGSVKVKKIIPKECHKMPAGEEQPV"
BASE COUNT      194 a 153 c 170 g 109 t
ORIGIN
Query Match      30.5%; Score 204; DB 9; Length 626;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
Qy 319 GTCCAGGAAACCCATGATGAAGAGCCTAAAGAGAGAAACCCACCTAAAGTCCGAA 378
Db 264 GTCCAGGAGCCAGTAATGGAGAGCCCAAAAAGAGAACCCAGAGCTGAAAGTCGGGA 323
Qy 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCGAGCTGAGATTCAGTCCCTGA 438
Db 324 TCCTACACCTGGCAGACAGAGAGAGATCAGATACAGCTGAGATCCCGTCGGGA 383
Qy 439 CCTGGAAGCCGATCTCCAGAGCTATGTCAGACAAAGACTGGGATGAGTGAAGGT-G 497
Db 384 CATGGAAGGTGATCTGCAAGAGCTGATCAGTCAAAACACCGGGGATAAATCTGGATTGG 443
Qy 498 GTACTGATGTCAAGGGAGATCTTACCAAAAGCAGAGCATTAAATGCCAGAGCAG 557
Db 444 GTTCCGGCTGAAGGTGAAGATATACCTTAAGAGGACACTGTAAATGCCAGAGCAG 503
Qy 558 GTGAAGGGAATCAGAGGTTTAAAGGAAGATAGCTGAAACACACAAACTGTTTTATA 617
Db 504 GTCAAGAGCAACACCAAGTTTAAATGAAGACAGCTGAAACACGCAAGCTGGTTTTATA 563
Qy 618 TTAGATATTTCCTTAAAGAGTCTTAATAAATTTT 653
Db 564 TTAGATATTTCCTTAAAGTCTTAATAAATTTT 599

RESULT 28
AX455517
LOCUS      AX455517      637 bp      DNA      linear      PAT 06-JUL-2002
DEFINITION Sequence 5 from Patent WO0218584.
ACCESSION  AX455517
VERSION     AX455517.1  GI:21714596
KEYWORDS    Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Pastan,I.H., Liu,X.F., Bera,T.K., Lee,B. and Eglund,K.A.
  TITLE     Xage-1, a gene expressed in multiple cancers, and uses thereof
  JOURNAL   Patent: WO 0218584-A 5 07-MAR-2002;
            THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)
FEATURES    Location/Qualifiers
  source    1..637
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /note="complete XAGE-1 sequence with untranslated 5' and
            3' ends"
  CDS       97..537
            /note="xage-1 p16"
            /codon_start=1
            /protein_id="CAD38240.1"
            /db_xref="GI:21714597"
            /translation="MLLWCFQCACSLGVFPSPVWGTTRSCPEPATRVPEVWLSP
            LLRHGHGTQNTHTASPRSPVMSPKKNQQLKVLHLSGROKIRIOLRSQCATWK
            VICKSCISQTPGINLDLGSVKVKKIIPKECHKMPAGEEQPV"

CDS
/tissue_type="testis"
86..526
/note="expressed in Ewing's sarcoma"
/codon_start=1
/product="XAGE-1"
/protein_id="XAGE-1"
/db_xref="GI:9885325"
/db_xref="GI:9885325"
/translation="MLLNCPPQCACSLGVFPSPVWGTTRSCPEPATRVPEVWLSP
LLRHGHGTQNTHTASPRSPVMSPKKNQQLKVLHLSGROKIRIOLRSQCATWK
VICKSCISQTPGINLDLGSVKVKKIIPKECHKMPAGEEQPV"
BASE COUNT      194 a 153 c 170 g 109 t
ORIGIN
Query Match      30.5%; Score 204; DB 9; Length 626;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
Qy 319 GTCCAGGAAACCCATGATGAAGAGCCTAAAGAGAGAAACCCACCTAAAGTCCGAA 378
Db 264 GTCCAGGAGCCAGTAATGGAGAGCCCAAAAAGAGAACCCAGAGCTGAAAGTCGGGA 323
Qy 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCGAGCTGAGATTCAGTCCCTGA 438
Db 324 TCCTACACCTGGCAGACAGAGAGAGATCAGATACAGCTGAGATCCCGTCGGGA 383
Qy 439 CCTGGAAGCCGATCTCCAGAGCTATGTCAGACAAAGACTGGGATGAGTGAAGGT-G 497
Db 384 CATGGAAGGTGATCTGCAAGAGCTGATCAGTCAAAACACCGGGGATAAATCTGGATTGG 443
Qy 498 GTACTGATGTCAAGGGAGATCTTACCAAAAGCAGAGCATTAAATGCCAGAGCAG 557
Db 444 GTTCCGGCTGAAGGTGAAGATATACCTTAAGAGGACACTGTAAATGCCAGAGCAG 503
Qy 558 GTGAAGGGAATCAGAGGTTTAAAGGAAGATAGCTGAAACACACAAACTGTTTTATA 617
Db 504 GTCAAGAGCAACACCAAGTTTAAATGAAGACAGCTGAAACACGCAAGCTGGTTTTATA 563
Qy 618 TTAGATATTTCCTTAAAGAGTCTTAATAAATTTT 653
Db 564 TTAGATATTTCCTTAAAGTCTTAATAAATTTT 599

RESULT 29
HSA318878
LOCUS      HSA318878      740 bp      mRNA      linear      PRI 14-JAN-2002
DEFINITION Homo sapiens mRNA for XAGE-1c protein.
ACCESSION  AJ318878
VERSION     AJ318878.1  GI:18157205
KEYWORDS    XAGE-1c gene,
            Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Zendman,A.J.W., van Kraats,A.A., Weidie,U.H., Ruiter,D.R. and Van
            Muljen,G.N.P.
  TITLE     Expression profile of members of the XAGE cancer/testis antigen
            family
  JOURNAL   Unpublished
  REFERENCE 2 (bases 1 to 740)
  AUTHORS   Zendman,A.J.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-AUG-2001) Zendman A.J.W., Department of Pathology,
            University Medical Center St Radboud, Geert Grooteplein Zuid 24, PO
            Box 9101 6500 HB Nijmegen, NETHERLANDS
FEATURES    Location/Qualifiers
  source    1..740
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /chromosome="X"
            /map="Xp11.21-Xp11.22"
            /tissue_type="Ewing sarcoma"
            /gene="XAGE-1c"
            /note="XAGE-1c"
            163..645

CDS
86..526
/note="expressed in Ewing's sarcoma"
/codon_start=1
/product="XAGE-1"
/protein_id="XAGE-1"
/db_xref="GI:9885325"
/db_xref="GI:9885325"
/translation="MLLNCPPQCACSLGVFPSPVWGTTRSCPEPATRVPEVWLSP
LLRHGHGTQNTHTASPRSPVMSPKKNQQLKVLHLSGROKIRIOLRSQCATWK
VICKSCISQTPGINLDLGSVKVKKIIPKECHKMPAGEEQPV"
BASE COUNT      194 a 153 c 170 g 109 t
ORIGIN
Query Match      30.5%; Score 204; DB 6; Length 637;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
Qy 319 GTCCAGGAAACCCATGATGAAGAGCCTAAAGAGAGAAACCCACCTAAAGTCCGAA 378
Db 275 GTCCAGGAGCCAGTAATGGAGAGCCCAAAAAGAGAACCCAGAGCTGAAAGTCGGGA 334
Qy 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCGAGCTGAGATTCAGTCCCTGA 438
Db 335 TCCTACACCTGGCAGACAGAGAGAGATCAGATACAGCTGAGATCCCGTCGGGA 394
Qy 439 CCTGGAAGCCGATCTCCAGAGCTATGTCAGACAAAGACTGGGATGAGTGAAGGT-G 497
Db 395 CATGGAAGGTGATCTGCAAGAGCTGATCAGTCAAAACACCGGGGATAAATCTGATTGG 454
Qy 498 GTACTGATGTCAAGGGAGATCTTACCAAAAGCAGAGCATTAAATGCCAGAGCAG 557
Db 455 GTTCCGGCTGAAGGTGAAGATATACCTTAAGAGGAAACACTGTAATAATGCCAGAGCAG 514
Qy 558 GTGAAGGGAATCAGAGGTTTAAAGGAAGATAGCTGAAACACACAAACTGTTTTATA 617
Db 515 GTGAAGAGCAACCAAGTTTAAATGAAGACAGCTGAAACACGCAAGCTGGTTTTATA 574
Qy 618 TTAGATATTTCCTTAAAGAGTCTTAATAAATTTT 653
Db 575 TTAGATATTTCCTTAAAGTCTTAATAAATTTT 610
```

/gene="XAGE-1c"
/codon_start=1
/product="XAGE-1c protein"
/protein_id="CAC82986.1"
/db_xref="GI:18157206"
/translation="WRCHAGPSCLVTAITREGGPRSGAQAALGCCWGPSPRSTW
NPDREFTWPTQPGRGHERHTOTQNHPTASPRSPVMBSPKKNNOLKVLGILHLSRQK
KIRQLRSQCATWKVICKSCLSQFPGINLDLGSQVKVLIIFKEHECKRPEAGEBPQPV
"
BASE COUNT 224 a 170 c 216 g 130 t
ORIGIN

Query Match 30.5%; Score 204; DB 9; Length 740;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
Qy 319 GTCCAGAGAACCCATGATGAGAGAGCTTAAGAAGAGAAACACCCCTAAAGTCGAA 378
Db 383 GTCCAGAGAGCCAGTAATGGAGAGCCCAAAAAGAGAACAGCAGCTGAAGTCGGGA 442
Qy 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGCTGAGCTGAGATTCAAGTGCCTGA 438
Db 443 TCCTACACCTGGCAGCAGAGAGAGATCAGATACAGTGAATCCAGTCGCGGA 502
Qy 439 CTGGAAGCCGATCTCCAGAGCTATGTCCAGCAAGAGCTGGGATGAGTGAAGGT-G 497
Db 503 CATGGAAGGTGATCTGCAAGAGCTGATCAGTCAAAACACCGGGGATAAATCTGGATTGG 562
Qy 498 GTACTGATCTCAAGGGAGATCTACCAAAAGCAGAGCATTAAATGCCAGAGCAG 557
Db 563 GTTCGGCTCAAGGTGAAGTAATACCTTAAGAGGACACTGTAAATGCCAGAGCAG 622
Qy 558 GTGAAGGGAATCACAGGTTTAAAGGAAGATAGCTGAAACACAACTGTTTTTATA 617
Db 623 GTGAGAGCAACCAAGTTTAAATGAGACAGCTGAAACACGCAAGCTGGTTTTATA 682
Qy 618 TTAGATATTACTTTAAAGAGCTCTTAATAATTTT 653
Db 683 TTAGATATTGACTTAAACTATCTCAATAAAGTTT 718

RESULT 30
BD109661
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
OS Homo sapiens (human)
PN JP 2002010789-A/1738
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EYE PI
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC Von Heijne matrix
CC score 4.6999980926514
CC seq LCALLSNFCPSST/VK
FH Key Location/Qualifiers
FT CDS 83..658

RESULT 31
AR273067
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
BASE COUNT 230 a 181 c 234 g 136 t
ORIGIN
Query Match 30.5%; Score 204; DB 6; Length 781;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
Qy 319 GTCCAGAGAACCCATGATGAGAGAGCTTAAGAAGAGAAACCCCTAAAGTCGAA 378
Db 422 GTCCAGAGAGCCAGTAATGGAGAGCCCAAAAAGAGAACCCAGCAGCTGAAGTCGGGA 481
Qy 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGCTGAGCTGAGATTCAAGTGCCTGA 438
Db 482 TCCTACACCTGGCAGCAGACAGAGAGATACGATACAGCTGAGATCCAGTGCCTGA 541
Qy 439 CCTGGAAGCCGATCTCCAGAGCTATGTCCAGCAAAAGACTGGGGATGGATGTGAAGGT-G 497

FT sig peptide 83..157.
Location/Qualifiers
source
1..767
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 222 a 193 c 187 g 165 t
ORIGIN

Query Match 30.5%; Score 204; DB 6; Length 767;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
Qy 319 GTCCAGAGAACCCATGATGAGAGAGCTTAAGAAGAGAAACCCCTAAAGTCGAA 378
Db 399 GTCCAGAGAGCCAGTAATGGAGAGCCCAAAAAGAGAACCCAGCAGCTGAAGTCGGGA 458
Qy 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGCTGAGCTGAGATTCAAGTGCCTGA 438
Db 459 TCCTACACCTGGCAGCAGACAGAGAGATCAGGATACAGCTGAGATCCAGTGCCTGA 518
Qy 439 CCTGGAAGCCGATCTCCAGAGCTATGTCCAGCAAAAGACTGGGGATGGATGTGAAGGT-G 497
Db 519 CATGGAAGGTGATCTCAAGAGCTGATCAGTCAAAACACCGGGGATAAATCTGATTGG 578
Qy 498 GTACTGATGTAAGGGAGAGATCTTACCAAAAGCAGAGCATTAAATGCCAGAGCAG 557
Db 579 GTTCGGCTCAAGGTGAAGTAATACCTTAAGAGAGAACACTGTAAATGCCAGAGCAG 638
Qy 558 GTGAAGGGAATCACAGGTTTAAAGGAAGATAGCTGAAACACAACTGTTTTTATA 617
Db 639 GTGAGAGCAACCAAGTTTAAATGAGAGCAAGCTGAAACACGCAAGCTGGTTTTATA 698
Qy 618 TTAGATATTACTTTAAAGAGCTCTTAATAATTTT 653
Db 699 TTAGATATTGACTTAAACTATCTCAATAAAGTTT 734

AR273067 781 bp DNA linear PAT 10-APR-2003
Sequence 808 from patent US 6504010.
AR273067
AR273067.1 GI:29704952
Unknown.
Unknown.
Unclassified.
1 (bases 1 to 781)
Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,
Carter, D., Retter, M.W., Mannion, J. and Fan, L.
Compositions and methods for the therapy and diagnosis of lung
cancer
Patent: US 6504010-A 808 07-JAN-2003;
Location/Qualifiers
1..781
/organism="unknown"
BASE COUNT 230 a 181 c 234 g 136 t
ORIGIN
Query Match 30.5%; Score 204; DB 6; Length 781;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
Qy 319 GTCCAGAGAACCCATGATGAGAGAGCTTAAGAAGAGAAACCCCTAAAGTCGAA 378
Db 422 GTCCAGAGAGCCAGTAATGGAGAGCCCAAAAAGAGAACCCAGCAGCTGAAGTCGGGA 481
Qy 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGCTGAGCTGAGATTCAAGTGCCTGA 438
Db 482 TCCTACACCTGGCAGCAGACAGAGAGATACGATACAGCTGAGATCCAGTGCCTGA 541
Qy 439 CCTGGAAGCCGATCTCCAGAGCTATGTCCAGCAAAAGACTGGGGATGGATGTGAAGGT-G 497

Db

542

CATGAGAGTGTCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGG

601

Qy

498

GTACTGATGTAAGGGAGATCTTACCAAAAGCAGAGACACTTTAAATGCCAGAGCAG

557

Db

602

GTTCGGCGCTCAAGTGAAGATAATACCTAAAGAGAACACTGTAAATGCCAGAGCAG

661

Qy

558

GTGAAGGGAAATCACAGGTTTAAAGGAAGATAAGCTGAAACACAACTGTTTTTATA

617

Db

662

GTGAAGAGCACACCAAGTTTAAATGAAGCAAGCTGAAACACGCAAGCTGGTTTTATA

721

Qy

618

TTAGATATTCTTAAAGAGCTTTAATAAATTTT

653

Db

722

TTAGATATTCTTAAAGAGCTTTAATAAATTTT

757

RESULT 32

AR276648

LOCUS

AR276648

Sequence 808 from patent US 6509448.

ACCESSION

AR276648

VERSION

AR276648.1

GI:29710295

KEYWORDS

Unknown.

SOURCE

Unknown.

ORGANISM

Unclassified.

REFERENCE

1 (bases 1 to 781)

AUTHORS

Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S., Carter,D., Retter,M.W., Mannion,J., Fan,L. and Wang,A.

TITLE

Compositions and methods for the therapy and diagnosis of lung cancer

JOURNAL

Patent: US 6509448-A 808 21-JAN-2003;

FEATURES

Location/Qualifiers

source

1..781

/organism="unknown"

BASE COUNT

230 a 181 c 234 g 136 t

ORIGIN

Query Match

30.5%; Score 204; DB 6; Length 781;

Best Local Similarity

77.4%; Pred. No. 1.2e-45;

Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

Qy

319

GTCCAGGAAACCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCTAAAGTCCGAA

378

Db

422

GTCCAGGAGCCAGTAAATGAGAGCCCCCAAAAGAGAACCCAGCTGAAAGTCCGGA

481

Qy

379

TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCGAGCTGAGATTCAAGTCCCTGA

438

Db

482

TCCTACACCTGGCAGCAGACAGAGAGATCAGGATACAGCTGAGATCCAGTCCGGA

541

Qy

439

CCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGATCGGGATGATGAGT-G

497

Db

542

CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGG

601

Qy

498

GTACTGATGTCAGGGGAGAGATTCTACCAAGCAGAGACCTTTAAATGCCAGAGCAG

557

Db

602

GTTCGGCGCTCAAGTGAAGATAATACCTAAAGAGGACACTGTAAATGCCAGAGCAG

661

Qy

558

GTGAAGGGAAATCACAGGTTTAAAGGAAGATAAGCTGAAACACAACTGTTTTTATA

617

Db

662

GTGAAGAGCAACCAACAGTTTAAATGAAGCAAGCTGAAACACGCAAGCTGGTTTTATA

721

Qy

618

TTAGATATTCTTAAAGAGCTTTAATAAATTTT

653

Db

722

TTAGATATTCTTAAAGAGCTTTAATAAATTTT

757

RESULT 34

AX368098

LOCUS

AX368098

Sequence 808 from Patent WO0204514.

ACCESSION

AX368098

VERSION

AX368098.1

GI:18856173

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1

AUTHORS

Wang,T., Watanabe,Y., Henderson,R.A., Johnson,J.C., Retter,M.W., Marnerakis,M., Carter,D., Fanger,G.R., Vedvick,T.S., Bangur,C.S., McNabb,A., Fanger,N., Switzer,A., McNeill,P.D. and Clapper,J.D.

TITLE

Compositions and methods for the therapy and diagnosis of lung cancer

JOURNAL

Patent: WO 0204514-A 808 17-JAN-2002;

FEATURES

Location/Qualifiers

source

1..781

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT

230 a 181 c 234 g 136 t

ORIGIN

Query Match

30.5%; Score 204; DB 6; Length 781;

Best Local Similarity

77.4%; Pred. No. 1.2e-45;

Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

Qy

319

GTCCAGGAAACCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCTAAAGTCCGAA

378

Db

422

GTCCAGGAGCCAGTAAATGAGAGCCCCCAAAAGAGAACCCAGCTGAAAGTCCGGA

481

Qy

379

TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCGAGCTGAGATTCAAGTCCCTGA

438

Db

482

TCCTACACCTGGCAGCAGACAGAGAGATCAGGATACAGCTGAGATCCAGTCCGGA

541

Qy

439

CCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGATGAGT-G

497

Db

542

CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGG

601

Qy

498

GTACTGATGTCAGGGGAGAGATTCTACCAAGCAGAGACCTTTAAATGCCAGAGCAG

557

Db

602

GTTCGGCGCTCAAGTGAAGATAATACCTAAAGAGGACACTGTAAATGCCAGAGCAG

661

Qy

558

GTGAAGGGAAATCACAGGTTTAAAGGAAGATAAGCTGAAACACAACTGTTTTTATA

617

Db

662

GTGAAGAGCAACCAACAGTTTAAATGAAGCAAGCTGAAACACGCAAGCTGGTTTTATA

721

Qy

618

TTAGATATTCTTAAAGAGCTTTAATAAATTTT

653

Db

722

TTAGATATTCTTAAAGAGCTTTAATAAATTTT

757

RESULT 33

AX063181

LOCUS

AX063181

Sequence 808 from Patent WO0100828.

ACCESSION

AX063181

VERSION

AX063181.1

GI:12541027

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1

AUTHORS

Wang,T., Watanabe,Y., Henderson,R.A., Johnson,J.C., Retter,M.W., Marnerakis,M., Carter,D., Fanger,G.R., Vedvick,T.S., Bangur,C.S., McNabb,A., Fanger,N., Switzer,A., McNeill,P.D. and Clapper,J.D.

TITLE

Compositions and methods for the therapy and diagnosis of lung cancer

JOURNAL

Patent: WO 0204514-A 808 17-JAN-2002;

FEATURES

Location/Qualifiers

source

1..781

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT

230 a 181 c 234 g 136 t

ORIGIN

Query Match

30.5%; Score 204; DB 6; Length 781;

Best Local Similarity

77.4%; Pred. No. 1.2e-45;

Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

Qy

319

GTCCAGGAAACCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCTAAAGTCCGAA

378

Db

422

GTCCAGGAGCCAGTAAATGAGAGCCCCCAAAAGAGAACCCAGCTGAAAGTCCGGA

481

Qy

379

TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCGAGCTGAGATTCAAGTCCCTGA

438

Db

482

TCCTACACCTGGCAGCAGACAGAGAGATCAGGATACAGCTGAGATCCAGTCCGGA

541

Qy

439

CCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGATGAGT-G

497

Db

542

CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGG

601

Qy

498

GTACTGATGTCAGGGGAGAGATTCTACCAAGCAGAGACCTTTAAATGCCAGAGCAG

557

Db

602

GTTCGGCGCTCAAGTGAAGATAATACCTAAAGAGGACACTGTAAATGCCAGAGCAG

661

Qy

558

GTGAAGGGAAATCACAGGTTTAAAGGAAGATAAGCTGAAACACAACTGTTTTTATA

617

Db

662

GTGAAGAGCAACCAACAGTTTAAATGAAGCAAGCTGAAACACGCAAGCTGGTTTTATA

721

Qy

618

TTAGATATTCTTAAAGAGCTTTAATAAATTTT

653

Db

722

TTAGATATTCTTAAAGAGCTTTAATAAATTTT

757

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1

AUTHORS

Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S., Carter,D., Retter,M.W. and Mannion,J.

TITLE

Compositions and methods for the therapy and diagnosis of lung cancer

JOURNAL

Patent: WO 0100828-A 808 04-JAN-2001;

FEATURES

Location/Qualifiers

source

1..781

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT

230 a 181 c 234 g 136 t

ORIGIN

Query Match

30.5%; Score 204; DB 6; Length 781;

Best Local Similarity

77.4%; Pred. No. 1.2e-45;

Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

Qy

319

GTCCAGGAAACCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCTAAAGTCCGAA

378

Db

422

GTCCAGGAGCCAGTAAATGAGAGCCCCCAAAAGAGAACCCAGCTGAAAGTCCGGA

481

Qy

379

TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCGAGCTGAGATTCAAGTCCCTGA

438

Db

482

TCCTACACCTGGCAGCAGACAGAGAGATCAGGATACAGCTGAGATCCAGTCCGGA

541

Qy

439

CCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGATCGGGATGATGAGT-G

497

Db

542

CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGG

601

Qy

498

GTACTGATGTCAGGGGAGAGATTCTACCAAGCAGAGACCTTTAAATGCCAGAGCAG

557

Db

602

GTTCGGCGCTCAAGTGAAGATAATACCTAAAGAGGACACTGTAAATGCCAGAGCAG

661

Qy

558

GTGAAGGGAAATCACAGGTTTAAAGGAAGATAAGCTGAAACACAACTGTTTTTATA

617

Db

662

GTGAAGAGCAACCAACAGTTTAAATGAAGCAAGCTGAAACACGCAAGCTGGTTTTATA

721

Qy

618

TTAGATATTCTTAAAGAGCTTTAATAAATTTT

653

Db

722

TTAGATATTCTTAAAGAGCTTTAATAAATTTT

757

RESULT 34

AX368098

LOCUS

AX368098

Sequence 808 from Patent WO0204514.

ACCESSION

AX368098

VERSION

AX368098.1

GI:18856173

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1

AUTHORS

Wang,T., Watanabe,Y., Henderson,R.A., Johnson,J.C., Retter,M.W., Marnerakis,M., Carter,D., Fanger,G.R., Vedvick,T.S., Bangur,C.S., McNabb,A., Fanger,N., Switzer,A., McNeill,P.D. and Clapper,J.D.

TITLE

Compositions and methods for the therapy and diagnosis of lung cancer

JOURNAL

Patent: WO 0204514-A 808 17-JAN-2002;

FEATURES

Location/Qualifiers

source

1..781

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT

230 a 181 c 234 g 136 t

ORIGIN

Query Match

30.5%; Score 204; DB 6; Length 781;

Best Local Similarity

77.4%; Pred. No. 1.2e-45;

Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

Qy

319

GTCCAGGAAACCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCTAAAGTCCGAA

378

Db

422

GTCCAGGAGCCAGTAAATGAGAGCCCCCAAAAGAGAACCCAGCTGAAAGTCCGGA

481

Qy

379

TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCGAGCTGAGATTCAAGTCCCTGA

438

Db

482

TCCTACACCTGGCAGCAGACAGAGAGATCAGGATACAGCTGAGATCCAGTCCGGA

541

Qy

439

CCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGATGAGT-G

497

Db

542

CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGG

601

Qy

498

GTACTGATGTCAGGGGAGAGATTCTACCAAGCAGAGACCTTTAAATGCCAGAGCAG

557

Db

602

GTTCGGCGCTCAAGTGAAGATAATACCTAAAGAGGACACTGTAAATGCCAGAGCAG

661

Qy

558

GTGAAGGGAAATCACAGGTTTAAAGGAAGATAAGCTGAAACACAACTGTTTTTATA

617

Db

662

GTGAAGAGCAACCAACAGTTTAAATGAAGCAAGCTGAAACACGCAAGCTGGTTTTATA

721

Qy

618

TTAGATATTCTTAAAGAGCTTTAATAAATTTT

653

Db

722

TTAGATATTCTTAAAGAGCTTTAATAAATTTT

757

```
BASE COUNT      230 a      181 c      234 g      136 t
ORIGIN
Query Match      30.5%; Score 204; DB 6; Length 781;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGAGACCCACTGATGAGAGCCCTTAAGAGAGAGAACACCCACTAAAGTCGAA 378
DB 422 GTCCAGGAGGCCAGTAATGGAGGCCCCCAAAAGAGAACCCAGCTGAAGTCGGGA 481
QY 379 TCTTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGA 438
DB 482 TCTTACACCTGGGCAGCAGCAGAGAGATCAGGATCAGCTGAGATCCCGCGGA 541
QY 439 CTTGGAGCCGATCTCCAGAGCTATGTCAGACCAAGCTGGGATGATGAGGT-G 497
DB 542 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGGATAAATCTGGATTGG 601
QY 498 GTACTGATGTCAAGGGGAGAGATTCTACCAAAAGCAGACACATTTTAAATGCCAGAGCAG 557
DB 602 GTTCCGGCTCAAGGTGAAGATAATACCTTAAGAGGAAACACATGTAAATGCCAGAGCAG 661
QY 558 GTGAGGGAATCAGAGGTTTAAAGGAAGATAGCTGAACCAACACAACTGTTTATA 617
DB 662 GTGAGAGCAGCACCAAGTTTAAATGAAGCAAGCTGAAACCAACCAAGCTGGTTTATA 721
QY 618 TTAGATATTCTTCTTAAAGAGCTCTTAATAATTTT 653
DB 722 TTAGATATTGACTTAAACTTCTCTCAATAAGTTT 757

RESULT 35
LOCUS      AR275665          528 bp      DNA      linear      PAT 10-APR-2003
DEFINITION Sequence 1 from patent US 6509172.
ACCESSION  AR275665
VERSION     AR275665.1 GI:29709167
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 528)
AUTHORS    De Backer, O., Van den Eynde, B. and Boon-Falleur, T.
TITLE      Isolated, truncated nucleic acid which are members of the gage, and
JOURNAL     Patent: US 6509172-A 1 21-JAN-2003;
FEATURES    Location/Qualifiers
            source
            1..528
            /organism="unknown"
BASE COUNT  158 a      112 c      151 g      107 t
ORIGIN
Query Match      27.3%; Score 182.8; DB 6; Length 528;
Best Local Similarity 67.6%; Pred. No. 9.6e-40;
Matches 288; Conservative 0; Mismatches 132; Indels 6; Gaps 2;
QY 217 GTGAATATGAGTTGGCGAGAGATCAATATAGCGCTAGCCCAAGAGAGTTTACA 276
DB 66 GTGAATATGAGTTGGCGAGAGATCGACTATCGCTAGCCCAAGAGAGTTTACA 125
QY 277 GCCTCTGAGCTGATTGGGGCTATGCTTACTGGCTCCCTTTGTGCCAGAAACCCACTGA 336
DB 126 GCCTCTGAAATGATTGGGGCTATGCGGCC---CGAGCAGTTTCAAGTGAAGTGAAC 182
QY 337 TGAAGAGCTTAAGAGAGAGAACCCCACTAAAGTCGGATCCTACACCTGATCAGAA 396
DB 183 AGCAACACCTGAAGAGAGAGAACCCAGCAACTCAACGTCAAGGCGGAGAGCTGATAGCA 242
QY 397 GAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAGAGCCGATCTCA 456
DB 243 GGGAGAGATG---AGGGAGCATCTGCAGGTCAAGGCGGAGAGCTGATAGCA 299

BASE COUNT      230 a      181 c      234 g      136 t
ORIGIN
Query Match      30.5%; Score 204; DB 6; Length 781;
Best Local Similarity 77.4%; Pred. No. 1.2e-45;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGAGACCCACTGATGAGAGCCCTTAAGAGAGAGAACACCCACTAAAGTCGAA 378
DB 422 GTCCAGGAGGCCAGTAATGGAGGCCCCCAAAAGAGAACCCAGCTGAAGTCGGGA 481
QY 379 TCTTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGA 438
DB 482 TCTTACACCTGGGCAGCAGCAGAGAGATCAGGATCAGCTGAGATCCCGCGGA 541
QY 439 CTTGGAGCCGATCTCCAGAGCTATGTCAGACCAAGCTGGGATGATGAGGT-G 497
DB 542 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGGATAAATCTGGATTGG 601
QY 498 GTACTGATGTCAAGGGGAGAGATTCTACCAAAAGCAGACACATTTTAAATGCCAGAGCAG 557
DB 602 GTTCCGGCTCAAGGTGAAGATAATACCTTAAGAGGAAACACATGTAAATGCCAGAGCAG 661
QY 558 GTGAGGGAATCAGAGGTTTAAAGGAAGATAGCTGAACCAACACAACTGTTTATA 617
DB 662 GTGAGAGCAGCACCAAGTTTAAATGAAGCAAGCTGAAACCAACCAAGCTGGTTTATA 721
QY 618 TTAGATATTCTTCTTAAAGAGCTCTTAATAATTTT 653
DB 722 TTAGATATTGACTTAAACTTCTCTCAATAAGTTT 757

RESULT 36
LOCUS      AF055473          528 bp      mRNA      linear      PRI 01-MAY-2000
DEFINITION Homo sapiens GAGE-8 mRNA, complete cds.
ACCESSION  AF055473
VERSION     AF055473.1 GI:3511022
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 528)
AUTHORS    Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 528)
            De Backer, O., Arden, K.C., Boretti, M., Vantomme, V., De Smet, C.,
            Crekay, S., Viars, C.S., De Plaen, E., Brasseur, F., Chomez, P., Van den
            Eynde, B., Boon, T. and van der Bruggen, P.
            Characterization of the GAGE genes that are expressed in various
            human cancers and in normal testis
            Cancer Res. 59 (13), 3157-3165 (1999)
JOURNAL     99323388
FEATURES    MEDLINE
            PUBMED 10397259
            2 (bases 1 to 528)
            De Backer, O.R.Y.
            Direct Submission
            Submitted (24-MAR-1998) Ludwig Institute for Cancer Research,
            Brussels Branch, 74 av. Hippocrate, Brussels B-1200, Belgium
            Location/Qualifiers
            source
            1..528
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            73..423
            /codon_start=1
            /product="GAGE-8"
            /protein_id="AAC33676.1"
            /db_xref="GI:3511023"
            /translation="MSWRGRSTYRPRVYEPPEMIGPMRPEQFDEVEPATPESGE
            PATQRQDFAAQSGEDGASAGGPKFADSGQHPQTCEDGDFDQEMDPFPE
            EVKTPESGEKQSQ"
BASE COUNT  158 a      112 c      151 g      107 t
ORIGIN
Query Match      27.3%; Score 182.8; DB 9; Length 528;
Best Local Similarity 67.6%; Pred. No. 9.6e-40;
Matches 288; Conservative 0; Mismatches 132; Indels 6; Gaps 2;
QY 217 GTGAATATGAGTTGGCGAGAGATCAATATAGCGCTAGCCCAAGAGAGATTTACA 276
DB 66 GTGAATATGAGTTGGCGAGAGATCGACTATCGCTAGCCCAAGAGAGCTACGTAGA 125
QY 277 GCCTCTGAGCTGATTGGGGCTATGCTTACTGGCTCCCTTTGTGCCAGAAACCCACTGA 336
DB 126 GCCTCTGAAATGATTGGGGCTATGCGGCC---CGAGCAGTTTCAAGTGAAGTGAAC 182
QY 337 TGAAGAGCTTAAGAGAGAGAACCCCACTAAAGTCGGATCCTACACCTGATCAGAA 396
DB 183 AGCAACACCTGAAGAGAGAGAACCCAGCAACTCAACGTCAAGGCGGAGAGCTGATAGCA 242
QY 397 GAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAGAGCCGATCTCA 456
DB 243 GGGAGAGATG---AGGGAGCATCTGCAGGTCAAGGCGGAGAGCTGATAGCA 299
```

QY 397 GAGAGAAGATGATCAGGCTGACAGTTCAGATTCAAGTGGCTGACCTGGGAAGCGGATCTCCA 456
 |||||
 Db 243 GGGAGAGGATG--AGGAGAGCATCTGCAGGTCAAGGGCGAAGCCTGAAGCTGATAGCCA 299
 |||||
 QY 457 GGAGCTATGTCAGACAAAGACTGGGATGGATGTGAAGTGTACTGTATGTCACAGGGAA 516
 |||||
 Db 300 GGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGCTCTGTATGGCAGGAT 359
 |||||
 QY 517 GATTCTACCAAGCAGAGACACTTTAAATGCGAGAGCAGGTGAAGGAAATCAAGGT 576
 |||||
 Db 360 GGACCCGCCAATCCAGAGGAGGTGAACCGCTGAAGAACTGAAGCAATCAAGTG 419
 |||||
 QY 577 TTAAGGAAGATGAAGCTGAAACACACAACTGTTTTATATAGATATTTACTTTAAA 636
 |||||
 Db 420 TTAAGGAAGACACGTTGAATGANGCAGGCTGCTCTATGTTGGAATTTGTTCAITAA 479
 |||||
 QY 637 GAGTCT 642
 |||||
 Db 480 AATTCT 485

RESULT 37
 HSIU19143 530 bp mRNA linear PRI 04-DEC-1995
 LOCUS Human GAGE-2 protein mRNA, complete cds.

DEFINITION
 U19143
 ACCESSION
 U19143.1 GI:914900

KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 530)
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

TITLE A new family of genes coding for an antigen recognized by
 autologous cytolytic T lymphocytes on a human melanoma

JOURNAL J. Exp. Med. 182 (3), 689-698 (1995)
 MEDLINE 95378788
 PUBMED 7544395

REFERENCE 2 (bases 1 to 530)
 AUTHORS Van Den Eynde, B.J.

TITLE Direct Submission
 JOURNAL Submitted (28-DEC-1994) Benoit J Van Den Eynde, Ludwig Institute
 For Cancer Research, 74 Avenue Hippocrate, BRUSSELS, 1200, BELGIUM

FEATURES
 source
 1..530
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /sex="female"
 /cell_line="M2-MEL.43"
 /tissue_type="melanoma"
 /dev_stage="adult"
 84..434
 /codon_start=1
 /product="GAGE-2 protein"
 /protein_id="AAA82745.1"
 /db_xref="GI:914901"

CDS
 1..530
 /translation="MSWRGRSTYRPRRYVEPPMIGMRPEQSFDEVPATPEEGE
 PATORDPAAQEGEDBGAGGQPKPEAHSRQGHFQTGCECEDPGDQEMDPNPE
 EVKTPERGEKQSC"

BASE COUNT 152 a 116 c 155 g 107 t
 ORIGIN

Query Match 27.1%; Score 181.2; DB 9; Length 530;
 Best Local Similarity 67.4%; Pred. No. 2.7e-39;
 Matches 287; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

QY 217 GTGAATATGATGGCGAGGAAGATCAACATATAGCGCTAGGCAAGAGAGTTTACA 276
 |||||
 Db 77 GTGAATATGATGGCGAGGAAGATCGACCTATCGCGCTAGACCAAGACGCTACGTAGA 136

QY 277 GCCTCCTCAGCTGATGGGGCTATGCTTACTGGCTCCCTTTGTCCAGAAACCCACTGA 336
 |||||
 Db 137 GCCTCCTGAAATGATTTGGGCTATGGGCT--CGAGCAGTTTCAGTGTGAAGTGAAC 193
 |||||
 QY 337 TGAAGAGCCTAAAGAGAGAAACCCACTAAAGTTCGGAATCTCTACACCTGATCAGAA 396
 |||||
 Db 194 AGCAACACCTTGAAGAGAGGGAACCACTCAAGTCAAGATCTCTGAGCTGCTCAGGA 253
 |||||
 QY 397 GAGAGAAGATGATCAGGGTGCAGCTCAGATTCAAGTGCCTGACCTGGAAAGCCGATCTCCA 456
 |||||
 Db 254 GGGAGAGGATG--AGGAGCATCTGCAGGTCAAGGGCCGAAGCTGAAGCTCATAGCCA 310
 |||||
 QY 457 GGAGCTATGTCAGACAAAGACTGGGATGGATGTGAAGTGTACTGTCTCAAGGGAA 516
 |||||
 Db 311 GGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGTGTGCTCTCTGATGGCAGAGAT 370
 |||||
 QY 517 GATTCTACCAAGCAGAGCACTTTAAATGCCAGAAAGCAGGTGAAGGAAATCAAGGT 576
 |||||
 Db 371 GGACCCGCCAATCCAGAGGAGGTGAACCGCTGAAGAAAGTGAAGCAATCAAGTG 430
 |||||
 QY 577 TTAAGGAAGATGAAGCTGAAACACACAACTGTTTTATATAGATATTTACTTTAAA 636
 |||||
 Db 431 TTAAGGAAGACACGTTGAATGANGCAGGCTGCTCTATGTTGGAATTTGTTCAITAA 490
 |||||
 QY 637 GAGTCT 642
 |||||
 Db 491 AATTCT 496

RESULT 38

AR028488

LOCUS

DEFINITION

AR028488

ACCESSION

AR028488.1

VERSION

GI:5940461

KEYWORDS

Unknown.

SOURCE

ORGANISM

Unclassified.

REFERENCE

1 (bases 1 to 538)

AUTHORS

van der Bruggen, P., van den Eynde, B., DeBacker, O. and

Boon-Falleur, T.

TITLE

Isolated peptides derived from the gage tumor rejection antigen

precursor and uses thereof

JOURNAL

Patent: US 5858689-A 14 12-JAN-1999;

FEATURES

Location/Qualifiers

1..538

source

/organism="unknown"

BASE COUNT 160 a 116 c 155 g 107 t

ORIGIN

Query Match 27.1%; Score 181.2; DB 6; Length 538;

Best Local Similarity 67.4%; Pred. No. 2.7e-39;

Matches 287; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

QY 217 GTGAATATGATGGCGAGGAAGATCAACATATAGCGCTAGGCAAGAGAGTTTACA 276

|||

Db 77 GTGAATATGATGGCGAGGAAGATCGACCTATCGCGCTAGACCAAGACGCTACGTAGA 136

|||

QY 277 GCCTCCTCAGCTGATGGGGCTATGCTTACTGGCTCCCTTTGTCCAGAAACCCACTGA 336

|||

Db 137 GCCTCCTGAAATGATTTGGGCTATGGGCT--CGAGCAGTTTCAGTGTGAAGTGAAC 193

|||

QY 337 TGAAGAGCCTAAAGAGAGAAACCCACTAAAGTTCGGAATCTCTACCTGATCAGAA 396

|||

Db 194 AGCAACACCTTGAAGAGAGGGAACCACTCAAGTCAAGATCTCTGAGCTGCTCAGGA 253

|||

QY 397 GAGAGAAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAAGCCGATCTCCA 456

|||

Db 254 GGGAGAGGATG--AGGAGCATCTGCAGGTCAAGGGCCGAAGCTGAAGCTCATAGCCA 310

|||

QY 457 GGAGCTATGTCAGACAAAGACTGGGATGGATGTGAAGTGGTACTGTATGTCAGAGGGAA 516

BC018052
LOCUS
DEFINITION
Homo sapiens, G antigen 8, clone MGC:26395 IMAGE:4812462, mRNA, complete cds.
ACCESSION
BC018052
VERSION
BC018052.1 GI:17390105
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 551)
Strausberg, R.
Direct Submission
Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Niklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILNL at: <http://image.llnl.gov>
Series: IRAK Plate: 32 Row: k Column: 1
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4503878.
Location/Qualifiers
1..551
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:26749"
/db_xref="taxon:9606"
/clone="MGC:26395 IMAGE:4812462"
/tissue_type="Brain, hippocampus"
/clone_lib="NIH MGC 95"
/lab_host="DH10B"
/note="Vector: pBluescript"
97..447
/codon_start=1
/product="G antigen 8"
/protein_id="AAH18052.1"
/db_xref="GI:17390106"
/translations="MSWRGRSTYRPRPRYPPEPMIGPMRPEQFSDEVEPATPEEGE
PATQRQDPAAQCEGEDEGASAGQGPKEADSOQHPTQCEGDGPDQEMDPNPE
EVTPEGEKQSQ"
BASE COUNT 164 a 118 c 160 g 109 t
ORIGIN
Query Match 27.1%; Score 181.2; DB 9; Length 551;
Best Local Similarity 67.4%; Pred. No. 2.7e-39;
Matches 287; Conservative 0; Mismatches 133; Indels 6; Gaps 2;
QY 217 GTGAAATATCAGTTGGCGAGAGATCAACATATAGCCCTAGCCCAAGAGAGTTTACA 276
DB 90 GTGAAATATGAGTTGGCGAGAGATCGACTATCGCCCTAGCAAGACGCTACGTAGA 149
QY GCCTCCTGAGCTGATGGGCTATGCTTACTGGCTCCCTTTGTCACAGAACCCACTGA 336
DB 150 GCCTCCTGAAATGATGGGCTATGCGGCC---CGAGCACTTCAGTGATGAAGTGAAC 206
QY 337 TGAAGCCCTAAGAGAGAAACCCACTAAAGTCGGAATCCCTACACCTGATCAGAA 396

Db 207 AGCAACACCTGAAGAAGGGGAACCACTCAACCTCAGGATCCTGCAGCTGCTCAGA 266
QY 397 GAGAGAAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCA 456
Db 267 GGGAGAGGATG---AGGGAGCATCTCAGGTCAAGGGCCGAAGCCTGAAGCTGATAGCA 323
QY 457 GGAGCTATGTCAGACAAAGACTGGGGATGGATGTGTGAAGTGGTACTGATGTCAAGGGGAA 516
Db 324 GGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAGATGCTCTGATGGCAGAGAT 383
QY 517 GATTTCACAAAAGCAGAGACCTTTTAAATGCCAAGCAGGTGAAGGGAATACAGGT 576
Db 384 GGACCCGCCAAATCCAGAGAGGTGAAACCGCTGAAGAGGTGAAGCAATACAGTG 443
QY 577 TTAAGAGAGATAGCTGAACCAACACAACTGTTTTTATATTAGATATTTTACTTTAAA 636
Db 444 TTAAGAGAGGACGCTTGAATATGTCAGCTGCTCTCTATGTTGGAATTTGTTCAATA 503
QY 637 GAGTCT 642
Db 504 AATTCT 509
RESULT 42
AR272358
LOCUS
DEFINITION
Sequence 70 from patent US 6504010.
ACCESSION
AR272358
VERSION
AR272358.1 GI:29704243
KEYWORDS
Unknown.
SOURCE
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 479)
Wang, T., Bangor, C.S., Lodes, M.J., Fanger, G.R., Vedwick, T.S., Carter, D., Retter, M.W., Mannion, J. and Fan, L.
Compositions and methods for the therapy and diagnosis of lung cancer
Patent: US 6504010-A 70 07-JAN-2003;
Location/Qualifiers
1..479
/organism="unknown"
BASE COUNT 163 a 107 c 125 g 84 t
ORIGIN
Query Match 26.9%; Score 180; DB 6; Length 479;
Best Local Similarity 74.0%; Pred. No. 5.8e-39;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGGACCCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCTAAAGTCCGAA 378
Db 103 GTCCAGAGGCCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCTGAAAGTCGGA 162
QY 379 TCCTACACCTGATCAGAAAGAGAGAGATGATCAGGGTGCAGCTGAGATTC----- 428
Db 163 TCCTACACCTGGGAGAGAGAGAGATGATGAGATGAGATGAGATGAGATGAGATGAGAT 222
QY 429 -----AAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGG 482
Db 223 GAAGGGAATCCGCGACATGGAAGTGTCTGCAAGAGCTGCTCATCAGTCAACACCGGG 282
QY 483 ATGATGTGAAGGT-GGTACTGATGTCAAGGGGAGATTTCTACAAAGCAGAGCACTTT 541
Db 283 ATAAATCTGGATTTGGGTTCCGGCTCAAGGTGAAGATAATACCTAAAGAGGAACTGT 342
QY 542 AAAATGCCAGAGAGAGGTGAAGGGAATCAAGGTTTAAAGGAGATGAAGCTGAACAC 601
Db 343 AAAATGCCAGAGAGAGGTGAAGGGAATCAAGGTTTAAAGGAGATGAAGCTGAACAC 402
QY 602 ACAAACTGTTTTTATATTAGATATTTTACTTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456

```
RESULT 43
AR275939
LOCUS AR275939 479 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 70 from patent US 6509448.
ACCESSION AR275939
VERSION AR275939.1 GI:29709584
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 479)
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,
Carter, D., Retter, M.W., Mannion, J., Fan, B. and Wang, A.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: US 6509448-A 70 21-JAN-2003;
FEATURES
LOCATION/Qualifiers
SOURCE 1. 479
/organism="unknown"
BASE COUNT 163 a 107 c 125 g 84 t
ORIGIN
Query Match 26.9%; Score 180; DB 6; Length 479;
Best Local Similarity 74.0%; Pred. No. 5.8e-39;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGGAAACCCACTGATGAAGAGCCTTAAGAGAAACCCACCCTAAAGTCGGAA 378
Db 103 GTCCAGGAGCCAGTAATGGAGAGCCCAAAAGAAAGAACCCAGCAGCTGAAAGTCGGGA 162
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCGAGCTGAGATTTC----- 428
Db 163 TCCTACACCTGGGCAGCAGACAGAGAGATCAGATACAGCTGAGATCCAGGTGCTGG 222
QY 429 -----AAGTGCTGACCTGGAGCGGATCTCCAGAGCTATGTCCAGAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGG 282
QY 483 ATGATGTGAAGT-CGTACTGATGTCAGGGGAGAGATTCTACCAAAAGCAGACACTTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGCGTCAGGTGAAGATAATACCTAAAGAGGAACTGT 342
QY 542 AAAATGCCAGAGCAGGTGAAGGAAATCACAGGTTTAAAGGAAGATAAGCTGAAACAAAC 601
Db 343 AAAATGCCAGAGCAGGTGAAGAGCAACCAACAAGTTTAAATGAAGCAAGCTGAAACAAAC 402
QY 602 ACAACTGTTTTATATTAGATATTTTACTTTAAAGAGCTTTAATAAATTTTG 655
Db 403 GCAAGCTGGTTTTATATTAGGATATTGACTTAAACTATCTCAATAAAGTTTG 456

RESULT 44
AX062443
LOCUS AX062443 479 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 70 from Patent WO0100828.
ACCESSION AX062443
VERSION AX062443.1 GI:12540318
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,
Carter, D., Retter, M.W. and Mannion, J.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0100828-A 70 04-JAN-2001;
FEATURES
LOCATION/Qualifiers
SOURCE 1. 479
/organism="Homo sapiens"
```

```
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 163 a 107 c 125 g 84 t
ORIGIN
Query Match 26.9%; Score 180; DB 6; Length 479;
Best Local Similarity 74.0%; Pred. No. 5.8e-39;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGGAAACCCACTGATGAAGAGCCTTAAGAGAAACCCACCCTAAAGTCGGAA 378
Db 103 GTCCAGGAGCCAGTAATGGAGAGCCCAAAAGAAAGAACCCAGCAGCTGAAAGTCGGGA 162
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCGAGCTGAGATTTC----- 428
Db 163 TCCTACACCTGGGCAGCAGACAGAGAGATCAGATACAGCTGAGATCCAGGTGCTGG 222
QY 429 -----AAGTGCTGACCTGGAGCGGATCTCCAGAGCTATGTCCAGAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGG 282
QY 483 ATGATGTGAAGT-CGTACTGATGTCAGGGGAGAGATTCTACCAAAAGCAGACACTTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGCGTCAGGTGAAGATAATACCTAAAGAGGAACTGT 342
QY 542 AAAATGCCAGAGCAGGTGAAGGAAATCACAGGTTTAAAGGAAGATAAGCTGAAACAAAC 601
Db 343 AAAATGCCAGAGCAGGTGAAGAGCAACCAACAAGTTTAAATGAAGCAAGCTGAAACAAAC 402
QY 602 ACAACTGTTTTATATTAGATATTTTACTTTAAAGAGCTTTAATAAATTTTG 655
Db 403 GCAAGCTGGTTTTATATTAGGATATTGACTTAAACTATCTCAATAAAGTTTG 456

RESULT 45
AX367360
LOCUS AX367360 479 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 70 from Patent WO0204514.
ACCESSION AX367360
VERSION AX367360.1 GI:18855464
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W.,
Marnerakis, M., Carter, D., Fanger, G.R., Vedvick, T.S., Bangur, C.S.,
McNabb, A., Fanger, N., Switzer, A., McNeill, P.D. and Clapper, J.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0204514-A 70 17-JAN-2002;
FEATURES
LOCATION/Qualifiers
SOURCE 1. 479
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 163 a 107 c 125 g 84 t
ORIGIN
Query Match 26.9%; Score 180; DB 6; Length 479;
Best Local Similarity 74.0%; Pred. No. 5.8e-39;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGGAAACCCACTGATGAAGAGCCTTAAGAGAAACCCACCCTAAAGTCGGAA 378
Db 103 GTCCAGGAGCCAGTAATGGAGAGCCCAAAAGAAAGAACCCAGCAGCTGAAAGTCGGGA 162
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCGAGCTGAGATTTC----- 428
Db 163 TCCTACACCTGGGCAGCAGACAGAGAGATCAGATACAGCTGAGATCCAGGTGCTGG 222
```

Qy 429 -----AAGTGCTGACCTGGAAGCGGATCTCCAGAGCTATGTGAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGCGACATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGG 282
Qy 483 ATGGATGTGAAGT-GGTACTGATGTCAAGGGGAAGATTCTACCAAAAGCAGAGCACTTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGCTCAAGGTGAAGATAATACCTAAAGAGGAACACTGT 342
Qy 542 AAAATGCCAGAACGAGTGAAGGAAATCACAGGTTTAAAGGAAGATAAGCTGAACAAC 601
Db 343 AAAATGCCAGAACGAGTGAAGGAAATCACAGGTTTAAATGAAGCAAGCTGAACAAC 402
Qy 602 ACAACTGTTTTATATTAGATATTTTACTTTAAGAGTCTTAATAAATTTTG 655
Db 403 GCAAGCTGTTTTATATTAGATATTTGACTTAACTATCTCAATAAAGTTTG 456

Search completed: January 18, 2004, 12:21:59
Job time : 1695.16 secs

OM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 17:24:54 ; Search time 124.524 Seconds
(without alignments)
14502.633 Million cell updates/sec

Title: US-10-051-835-1
Perfect score: 669
Sequence: 1 cacacacgacgacgactt.....tttttggcagtcgatctc 669

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :			
N_Geneseq_19Jun03.*			
1:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*		
2:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*		
3:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*		
4:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*		
5:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*		
6:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*		
7:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*		
8:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*		
9:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*		
10:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*		
11:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*		
12:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*		
13:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*		
14:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*		
15:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*		
16:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*		
17:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*		
18:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*		
19:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*		
20:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*		
21:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*		
22:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*		
23:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*		
24:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*		
25:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*		
SUMMARIES			
Result No.	Score	Query Match length	ID Description
1	581.8	87.0	673 22 AAK51918 Human polynucleoti
2	578.6	86.5	642 24 ABA93876 Human G protein-co
3	558	83.4	756 25 ABT15738 Human cancer/testi
4	549	82.1	580 22 AAF59637 Human cell cycle a
5	507.8	75.9	611 22 AAD14983 Human NOV4 DNA H
6	469.2	70.1	665 22 AAK52302 Human polynucleoti
7	317	47.4	659 22 AAI58744 Human polynucleoti
8	296.2	44.3	532 22 AAI60530 Human polynucleoti

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

9	290.2	43.4	509	25	ABX77605	Differentially exp
10	288.6	43.1	520	25	ABT15737	Human cancer/testi
11	280.8	42.0	503	23	AA569484	DNA encoding novel
12	242.4	36.2	661	23	AA569486	DNA encoding novel
13	215	32.1	320	22	AAH83260	Human ovarian tumo
14	214	32.0	214	22	AA524637	Human ovarian PCR-
15	204	30.5	399	22	AAF68151	Human lung tumour
16	204	30.5	399	24	ABK38062	cDNA encoding Clon
17	204	30.5	399	25	ACA10391	Human lung cancer-
18	204	30.5	399	25	ABX99342	Lung cancer therap
19	204	30.5	457	22	AAF68851	Human lung tumour
20	204	30.5	457	24	ABK38762	cDNA encoding L552
21	204	30.5	457	25	ACA11091	Human lung cancer-
22	204	30.5	457	25	ACA02277	Lung cancer therap
23	204	30.5	463	24	AA92217	Melanoma metastasi
24	204	30.5	626	25	ACS1056	Human bladder canc
25	204	30.5	626	25	ABX76178	Lung cancer-associ
26	204	30.5	637	24	ABK50912	XAGE-1 gene encodi
27	204	30.5	762	22	AAH64751	Human lung tumour
28	204	30.5	781	22	AAF68861	Human lung tumour
29	204	30.5	781	24	ABK38772	cDNA encoding L552
30	204	30.5	781	25	ACA11101	Human lung cancer-
31	204	30.5	781	25	ACA02287	Lung cancer therap
32	195	29.1	822	22	AA122641	Human breast canc
33	191.2	28.6	618	22	AA113774	Human breast canc
34	183.8	27.5	750	23	AA591235	DNA encoding novel
35	181.2	27.1	530	25	AB220463	GAGE-2 full length
36	181.2	27.1	535	19	AAV18717	cDNA encoding GAGE
37	181.2	27.1	538	20	AA90519	GAGE-2 tumour reje
38	180	26.9	479	22	AAF68152	Human lung tumour
39	180	26.9	479	24	ABK38063	cDNA encoding clon
40	180	26.9	479	25	ACA10392	Human lung cancer-
41	180	26.9	479	25	ABX99343	Lung cancer therap
42	179.6	26.8	1245	22	AA560826	Human cancer agent
43	178	26.6	461	22	AAF68852	Human lung tumour
44	178	26.6	461	24	ABK38763	cDNA encoding L552
45	178	26.6	461	25	ACA11092	Human lung cancer-
46	178	26.6	461	25	ACA02278	Lung cancer therap
47	169.8	25.4	527	25	AA576236	Human bladder canc
48	169.8	25.4	527	25	ABX76236	Lung cancer-associ
49	169.8	25.4	528	24	ABL66323	Lung cancer relate
50	169.8	25.4	532	19	AAV18720	cDNA encoding GAGE
51	169.8	25.4	532	20	AA90522	GAGE-5 tumour reje
52	169.8	25.4	539	19	AAV18721	cDNA encoding GAGE
53	169.8	25.4	540	20	AA90521	GAGE-4 tumour reje
54	169.8	25.4	540	20	AA90523	GAGE-6 tumour reje
55	163.8	24.5	560	19	AAV18718	cDNA encoding GAGE
56	162.2	24.2	560	20	AA90520	GAGE-3 tumour reje
57	158.8	23.7	541	19	AAV18719	cDNA encoding GAGE
58	158.8	23.7	486	24	ABK38066	DNA encoding recom
59	158.6	23.7	486	24	ABK38068	DNA encoding recom
60	158.6	23.7	486	25	ACA03321	Lung cancer therap
61	158.6	23.7	486	25	ACA03323	Lung cancer therap
62	158.6	23.7	507	24	ABK39799	DNA encoding recom
63	158.6	23.7	507	25	ACA12128	human His tagged l
64	158.6	23.7	507	25	ACA03314	Lung cancer therap
65	155.2	23.2	288	24	ABK39761	cDNA encoding lung
66	155.2	23.2	288	25	ACA12090	Human lung cancer
67	155.2	23.2	288	25	ACA03276	Lung cancer therap
68	148.6	22.2	656	17	AA297402	Human prostate can
69	144.2	21.6	646	17	AA297402	GAGE tumour reject
70	144.2	21.6	646	19	AAV05540	cDNA encoding GAGE
71	144.2	21.6	646	20	AA90518	GAGE tumour reject
72	142.8	21.3	1024	21	AA297217	Human prostate can
73	142.8	21.3	648	16	AAQ81405	GAGE tumour reject
74	141	21.1	650	21	AA559322	cDNA encoding a ca
75	140.8	21.0	1024	21	AA559322	Human prostate can
76	138.2	20.7	516	25	ABT16028	NOVX related polyn
77	137	20.5	430	21	AAAC02129	Human secreted pro
78	134.6	20.1	334	25	ABT15743	Human cancer/testi
79	124.2	18.6	528	25	ABT15728	Human cancer/testi
80	124.2	18.6	735	25	ACS1026	Human bladder canc
81	124.2	18.6	763	25	ABX76235	Lung cancer-associ

c

82 123.2 18.4 475 22 AAD14981 Human NOV2 DNA. H
 83 123.2 18.4 538 25 AAT15736 Human cancer/testi
 84 122.4 18.3 124 21 AAC110466 Human secreted pro
 85 122.2 18.3 300 20 AAX55997 Human cDNA clone F
 86 118.8 17.8 215 25 ABZ78639 Tumour suppression
 87 118.8 17.8 215 25 ABZ09186 Human oligonucleot
 88 116.8 17.5 1051 22 AAD14982 Human NOV3 DNA. H
 89 114 17.0 530 22 AAS60104 Human cancer agent
 90 110.4 16.5 365 22 AAS60496 Human cancer agent
 91 110.4 16.5 2168 25 ABZ73737 Secreted protein g
 92 108.8 16.3 2182 21 AAC59112 Human secreted pro
 93 108.8 16.3 2182 25 ABZ73297 Secreted protein-e
 94 108.2 16.2 121724 24 ABQ88143 Human osteoblast d
 95 102.8 15.4 348 22 AAD14997 PCR2.1-83350421-S7
 96 100.8 15.1 201 21 AAD43613 Human secreted exp
 97 99.8 14.9 443 24 ABK64575 Human benign prost
 98 99.2 14.8 346 24 ABK64338 Human benign prost
 99 99.2 14.8 346 24 ABL69505 Prostate cancer re
 100 99.2 14.8 420 21 AAA06691 Human immunogenic

ALIGNMENTS

RESULT 1
 AAK51918
 ID AAK51918 standard; cDNA; 673 BP.

AC AAK51918;

DT 06-NOV-2001 (first entry)

DE Human polynucleotide SEQ ID NO 463.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KN tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.

OS Homo sapiens.

XX WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US04098.

XX 03-FEB-2000; 2000US-0496914.

XX 27-APR-2000; 2000US-0560875.

XX 20-JUN-2000; 2000US-0598075.

XX 19-JUL-2000; 2000US-0620325.

XX 01-SEP-2000; 2000US-0654936.

XX 15-SEP-2000; 2000US-0683561.

XX 20-OCT-2000; 2000US-0693325.

XX 30-NOV-2000; 2000US-0728422.

XX (HYSE-) HYSEQ INC.

XX WPI; 2001-476283/51.

XX P-PSDB; AAM78785.

XX Claim 1; Page 1680-1681; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.

XX Sequence 673 BP; 203 A; 145 C; 170 G; 155 T; 0 other;

Query Match 87.0%; Score 581.8; DB 22; Length 673;

Best Local Similarity 95.7%; Pred. No. 2.5e-166;

Matches 617; Conservative 0; Mismatches 7; Indels 21; Gaps 1;

Qy 10 GGCACGAGCTTCAGTGTGCATGTTCTTGGACACCTGCGTCAGTGTGCATGTTCTACTGGG 59

Db 26 GGCACGAGCTTCAGTGTGCATGTTCTTGGACACCTGCGTCAGTGTGCATGTTCTACTGGG 85

Qy 70 CATCTTCCCTTCGACCGCTTTGCCCACGCTGGTCCACGCTGGGAGCTGTGAGAGTGTGAG 129

Db 86 CATCTTCCCTTCGACCGCTTTGCCCACGCTGGTCCACGCTGGGAGCTGTGAGAGTGTGAG 145

Qy 130 GGGCACGCTTCGACCGCTTCGACTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 199

Db 146 GGGCACGCTTCGACCGCTTCGACTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 205

Qy 190 AGGCGAGCTCTCCGAGGAAGTGAATAGTGAATATGAGTTGGCGAGGAGATCAACATA 249

Db 206 AGGCGAGCTCTCCGAGGAAGTGAATAGTGAATATGAGTTGGCGAGGAGATCAACATA 265

Qy 250 TAGCCCTAGCCCAAGAAAGTTTACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 309

Db 266 TAGCCCTAGCCCAAGAAAGTTTACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 320

Qy 310 CTCCCTCTTCTCCAGCAACCCACTGATGAAGAGCTTAAAGAGAGAGAAACCCACCTAA 369

Db 321 -----GACCCACTGATGAAGAGCTTAAAGAGAGAGAAACCCACCTAA 364

Qy 370 AAGTCGGAATCCTACACCTGATCAGAAAGAGAGATGATCAGGCTGCAGCTGAGATTCA 429

Db 365 AAGTCGGAATCCTACACCTGATCAGAAAGAGAGATGATCAGGCTGCAGCTGAGATTCA 424

Qy 430 AGTCGCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGCAAAAGACTGGGATGATG 489

Db 425 AGTCGCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGCAAAAGACTGGGATGATG 484

Qy 490 TGAAGGTGCTACTGATGTCAAGGGGAAAGATTCTACCAAAAAGCAGAGCACTTTAAATGCC 549

Db 485 TGAAGGTGCTACTGATGTCAAGGGGAAAGATTCTACCAAAAAGCAGAGCACTTTAAATGCC 544

Qy 550 AGAGCAGCTGAAGGGAATCAAGGTTTAAAGGAAGATAGCTGAACACACAAACTG 609

Db 545 AGAGCAGCTGAAGGGAATCAAGGTTTAAAGGAAGATAGCTGAACACACAAACTG 604

Qy 610 TTTTATATATAGATATTTTACTTTTAAAGAGCTCTTAAATAAATTTT 654

Db 605 TTTTATATATAGATATTTTACTTTTAAATAAATTTT 649

RESULT 2

ABA93876

ID ABA93876 standard; cDNA; 642 BP.

XX AC ABA93876;

XX DT 07-MAY-2002 (first entry)

XX Human G protein-coupled receptor NOV4 encoding cDNA SEQ ID NO:9.

XX DE

KW Human; G protein-coupled receptor; receptor; GPCR; NOVX; NOVX; cardiant;
 KW antiarteriosclerotic; cytostatic; immunomodulatory; anorectic; antiviral;
 KW antiinflammatory; coagulant; gene therapy; cardiomyopathy; hypertension;
 KW atherosclerosis; valve disease; obesity; prostate cancer; adenocarcinoma;
 KW uterine cancer; fertility; haemophilia; immunodeficiency; AIDS;
 KW graft versus host disease; Crohn's disease; multiple sclerosis; ss.
 XX
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH CDS 207..542
 FT /*tag= a
 FT /product= "NOV4 protein"

XX WO200200691-A2.

XX 03-JAN-2002.

XX 27-JUN-2001; 2001WO-US20510.

XX 27-JUN-2000; 2000US-214759P.

XX 31-OCT-2000; 2000US-244546P.

XX 13-NOV-2000; 2000US-248153P.

XX 11-JAN-2001; 2001US-261014P.

XX 22-JAN-2001; 2001US-263215P.

XX (CURA-) CURAGEN CORP.

XX Vernet CAM, Tchernev V, Putturajan M, Malyankar UM, Gusev V;
 PI Hermann JJ, MacDougall JR, Rastelli L, Zhong H, Spytek KA;
 PI Shency S, Gerlach VL, Gangolli EA, Stone DJ, Smithson G;
 XX

DR WPI: 2002-090517/12.

DR P-PSDB; ABB05753.

XX Nucleic acids encoding polypeptides, designated NOVX polypeptides,
 PT useful for treating a syndrome associated with a NOVX-associated
 PT disorder, e.g. cardiomyopathy, atherosclerosis, hypertension, valve
 PT diseases, obesity, and prostate cancer -

XX Claim 8; Page 43; 215pp; English.

CC The present invention describes novel human G protein-coupled receptors
 CC (GPCRs), designated NOVX proteins (i.e. a NOV protein where X is one
 CC of: 1a; 1b; 2; 3; 4; 5; 6a; 6a2; 6a3; 6b; 6c; 6d; 7a; 7b; and 7c).
 CC NOVX proteins and polynucleotides have cardiant, antiarteriosclerotic,
 CC cytostatic, immunomodulatory, anorectic, antiviral, antiinflammatory
 CC and coagulant activities, and can be used in gene therapy. The NOVX
 CC proteins and polynucleotides are useful for treating cardiomyopathy,
 CC atherosclerosis, hypertension, valve diseases, obesity, prostate cancer,
 CC adenocarcinoma, uterus cancer, fertility, haemophilia, Crohn's disease,
 CC immunodeficiencies, AIDS, graft versus host disease or multiple
 CC sclerosis. The present sequence represents a human NOVX protein encoding
 CC cDNA sequence from the present invention.

XX Sequence 642 BP; 197 A; 136 C; 160 G; 149 T; 0 other;

XX Query Match 86.5%; Score 578.6; DB 24; Length 642;
 XX Best Local Similarity 96.1%; Pred. No. 2.3e-165;
 XX Matches 612; Conservative 0; Mismatches 4; Indels 21; Gaps 1;

QY 18 CTTTCAGTGTGCATGTCCTTGACACACCTGCTCCTCAGTGTGCATGTCATGGGCATCTTCC 77
 DB 1 CTTTCAGTGTGCATGTCCTTGACACACCTGCTCCTCAGTGTGCATGTCATGGGCATCTTCC 60

QY 78 CTTTCAGTGTGCATGTCCTTGACACACCTGCTCCTCAGTGTGCATGTCATGGGCACGT 137

DB 61 CTTTCAGTGTGCATGTCCTTGACACACCTGCTCCTCAGTGTGCATGTCATGGGCACGT 120

QY 138 TCCAGCCGTCTGACATCTTCTCTCCTACTGAGACGACGCTATAGGTCCGAGGCCAGT 197

DB 121 TCCAGCCGTCTGACATCTTCTCTCCTACTGAGACGACGCTATAGGTCCGAGGCCAGT 180

QY 198 CTTCCAGGAACTGAAATAGTGAATATGAGTTGGCGAGGAGATCAACATATAGGCCTA 257
 DB 181 CTTCCAGGAACTGAAATAGTGAATATGAGTTGGCGAGGAGATCAACATATAGGCCTA 240
 QY 258 GGCCTAAGAAAGTTTACAGCCTCTCTGAGCTGATTTGGGGCTATGCTTACTGCTCCCTT 317
 DB 241 GGCCTAAGAAAGTTTACAGCCTCTCTGAGCTGATTTGGGGCTATGCTTACTGCTCCCTT 287
 QY 318 TGTCCAGGAACTGATGATGAGGCTTAAAGAGAGAAACCCACCTAAAGTCCGA 377
 DB 288 -----GAACTGATGAGGCTTAAAGAGAGAAACCCACCTAAAGTCCGA 339
 QY 378 ATCTACCTGATCAGAAAGAGAGATGATCAGGTGTCAGCTGAGATTCAAGTCCCTG 437
 DB 340 ATCTACCTGATCAGAAAGAGAGATGATCAGGTGTCAGCTGAGATTCAAGTCCCTG 399
 QY 438 ACCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGGATGTGAAGTG 497
 DB 400 ACCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGGATGTGAAGTG 459
 QY 498 GTACTGATCTCAAGGGGAAAGATTCTACCAAGAGAGAGACATTTAAATGCCAGAGCAG 557
 DB 460 GTACTGATCTCAAGGGGAAAGATTCTACCAAGAGAGAGACATTTAAATGCCAGAGCAG 519
 QY 558 GTGAGGGGAAATCACAGGTTTAAAGGAAGATAGCTGAAACCAACAACTGTTTTATA 617
 DB 520 GTGAGGGGAAATCACAGGTTTAAAGGAAGATAGCTGAAACCAACAACTGTTTTATA 579
 QY 618 TTAGATATTTTACTTTTAAAGAGCTTTTAAATTTT 654
 DB 580 TTAGATATTTTACTTTTAAAGAGCTTTTAAATTTT 616

RESULT 3

ABT15738

ID ABT15738 standard; DNA; 756 BP.

XX AC ABT15738;

XX DT 28-MAR-2003 (first entry)

XX DE Human cancer/testis antigen coding sequence - SEQ ID No 40.

XX KW Human; gene; ds; gene therapy; vaccine; cancer; cancer/testis antigen;
 XX CT antigen.

XX OS Homo sapiens.

XX PN WO200278526-A2.

XX PD 10-OCT-2002.

XX XX 29-MAR-2002; 2002WO-US09808.

XX XX 30-MAR-2001; 2001US-280718P.

XX XX 20-APR-2001; 2001US-285154P.

XX XX 05-OCT-2001; 2001US-327432P.

XX XX 22-JAN-2002; 2002US-0054683.

XX XX (LUDW-) LUDWIG INST CANCER RES.

XX XX (CORR) CORNELL RES FOUND INC.

XX XX Old LJ, Scanlan MJ, Chen Y;

XX XX WPI; 2003-040608/03.

XX XX P-PSDB; ABJ19257.

XX XX Diagnosing cancer comprises contacting a biological sample isolated
 PT from a subject with an agent that specifically binds to a nucleic acid
 PT molecule, its expression product or fragment or an antibody that binds
 PT to the product or fragment -

XX PS Claim 8; Page 146; 155pp; English.

XX The invention comprises a method for diagnosing cancer, the method
CC involves detecting the DNA or protein sequences of human cancer/testis
CC (CT) antigens that are disclosed in the invention. The method of the
CC invention is useful for detecting/diagnosing, treating and monitoring a
CC cancer or condition characterised by the expression of a human CT
CC antigen. The present DNA sequence encodes a human CT antigen of the
CC invention.
SQ Sequence 756 BP; 223 A; 150 C; 216 G; 167 T; 0 other;

Query Match 83.4%; Score 558; DB 25; Length 756;
Best Local Similarity 95.8%; Pred. No. 4.5e-159;
Matches 592; Conservative 0; Mismatches 5; Indels 21; Gaps 1;

QY 37 TGGACACCTGCTCAGTGTGATGTTTCACTGGGCGATCTTCCCTTCGACCCCTTTGCCAC 96
Db 41 TGGACACCTGCTCAGTGTGATGTTTCACTGGGCGATCTTCCCTTCGACCCCTTTGCCAC 100

QY 97 GTGGTCACCGCTGGGAGCTGTGAGAGTGTGAGGGGACGTTCCAGCGTCTGGACTCTT 156
Db 101 GTGGTCACCGCTGGGAGCTGTGAGAGTGTGAGGGGACGTTCCAGCGTCTGGACTCTT 160

QY 157 TCTCTCTACTGAGACGCGCTATAGGTCCGAGCCAGTCTCCAGGAACTGAATA 216
Db 161 TCTCTCTACTGAGACGCGCTATAGGTCCGAGCCAGTCTCCAGGAACTGAATA 220

QY 217 GTCAATATGAGTGGCGAGGAGATCAACATATAGGCTAGGCCAAGAGAGTTTACA 276
Db 221 GTCAATATGAGTGGCGAGGAGATCAACATATAGGCTAGGCCAAGAGAGTTTACA 280

QY 277 GCCTCTGAGCTGATTGGGGCTATGCTTACTGCTCCCTTTTGTCCAGGAACCCACTGA 336
Db 281 GCCTCTGAGCTGATTGGGGCTATGCTT-----GAACCCACTGA 319

QY 337 TGAAGACCTTAAAGAGAGAAACCCACTAAAGTCCGAACTTACACCTGATCAGAA 396
Db 320 TGAAGACCTTAAAGAGAGAAACCCACTAAAGTCCGAACTTACACCTGATCAGAA 379

QY 397 GAGAGAAGATGATCAGGGTGCAGGTGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCA 456
Db 380 GAGAGAAGATGATCAGGGTGCAGGTGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCA 439

QY 457 GGAGCTATGTCAGCAAGACTGGGATGATGTGAGGTGTACTGATCTCAAGGGAA 516
Db 440 GGAGCTATGTCAGCAAGACTGGGATGATGTGAGGTGTACTGATCTCAAGGGAA 499

QY 517 GATTCTACCAAGACAGACACTTTAAATGCCAGAGAGAGTGAAGGGAAATCAAGGT 576
Db 500 GATTCTACCAAGACAGACACTTTAAATGCCAGAGAGAGTGAAGGGAAATCAAGGT 559

QY 577 TTAAGGAAGATAGCTGAACACAACTGTTTATATTAGATATTTTACTTTAA 636
Db 560 TTAAGGAAGATAGCTGAACACAACTGTTTATATTAGATATTTTACTTTAA 619

QY 637 GAGTCTTAATAAATTTT 654
Db 620 ATAATCTTAATAAATTTT 637

RESULT 4

AAF59637

ID AAF59637 standard; cDNA; 580 BP.

XX AC

XX XX

XX XX

DT 24-APR-2001 (first entry)

XX Human cell cycle and proliferation protein CCYPR-48 cDNA, SEQ ID NO:102.

DE XX

XX KW

KW antagonist; gene therapy; immune disorder;

KW transgenic animal disease model; immune disorder;

KW

KW developmental disorder; cell signalling disorder;
KW cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
KW arteriosclerosis; asthma; allergy; diabetes mellitus;
KW menstrual cycle disorder; bacterial infection; ss.
XX Homo sapiens.
XX WO200107471-A2.
XX 01-FEB-2001.
XX 21-JUL-2000; 2000WO-US19948.
XX 21-JUL-1999; 99US-0145075.
XX 08-SEP-1999; 99US-0153129.
XX 10-NOV-1999; 99US-0164647.
XX (INCY-) INCYTE GENOMICS INC.
XX Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;
PI Azimzai Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;
XX WPI: 2001-112727/12.
DR P-FSDB; AAB60500.
XX Human cell cycle and proliferation proteins and polynucleotides are
PT used to treat, diagnose and prevent immune, developmental and cell
PT signalling disorders and cell proliferative disorders including cancer -
XX Claim 5; Page 200; 205pp; English.
XX Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human
CC cell cycle and proliferation proteins (CCYPR), AAB60453-AAB60506.
CC CCYPR and agonists of CCYPR are used to treat diseases or conditions
CC associated with decreased expression of functional CCYPR, while CCYPR
CC antagonists are used to treat diseases or conditions associated with
CC overexpression of functional CCYPR. Monoclonal or polyclonal antibodies
CC to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or
CC radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
CC compounds e.g., antibodies, oligonucleotides and proteins (receptors)
CC that specifically bind to CCYPR, and in drug screening methods to
CC identify compounds that modulate the activity of CCYPR. CCYPR
CC nucleotides can be used to generate transgenic animal models of human
CC disease, and can be used in gene therapy in target cells with genetic
CC abnormalities with respect to the expression of CCYPR for the
CC treatment or prevention of a disorder associated with CCYPR.
CC Diseases which can be diagnosed, treated and prevented using CCYPR
CC proteins, nucleic acids, agonists or antagonists include immune,
CC developmental and cell signalling disorders, and cell proliferative
CC disorders including cancer. Specific examples of these disorders
CC include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
CC diabetes mellitus, disorders of the menstrual cycle and infections
CC caused by bacteria.
XX Sequence 580 BP; 167 A; 135 C; 160 G; 118 T; 0 other;
SQ

Query Match 82.1%; Score 549; DB 22; Length 580;

Best Local Similarity 96.5%; Pred. No. 2.1e-156;

Matches 580; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

QY 1 CACAACGAGGACCGACCTTTCAGTGTGATGTTCTTGGACACCTGCTCAGTGTGATG 60

Db 1 CACAACGAGGACCGACCTTTCAGTGTGATGTTCTTGGACACCTGCTCAGTGTGATG 60

QY 61 TTCACCTGGGCGATCTTCCCTTCGACCCCTTTGCCACCGTGTGACCGTGGGAGCTGTGA 120

Db 61 TTCACCTGGGCGATCTTCCCTTCGACCCCTTTGCCACCGTGTGACCGTGGGAGCTGTGA 120

QY 121 GAGTGTGAGGGGACGTTCCAGCCGCTCTGGACTCTTCTCTCTCTCTCTCTCTCTCTCT 180

Db 121 GAGTGTGAGGGGACGTTCCAGCCGCTCTGGACTCTTCTCTCTCTCTCTCTCTCTCTCT 180

QY 181 TAGGTCCGCGAGGCCAGTCTCTCCAGGAAGTGAATATGTAATATGAGTGTGGCGAGGAAG 240


```

Db 191 TAGGTCCGAGCCAGTCTCCAGAACTGAATAGTGAATATGATTGGCAGGAAG 240
Qy 241 ATCAACATATAGGCTAGGCCAAGAAAGTTTACAGCCTCTCAGCTGTGATTTGGGGCTAT 300
Db 241 ATCAACATATAGGCTAGGCCAAGAAAGTTTACAGCCTCTCAGCTGTGATTTGGGGCTAT 300
Qy 301 GCTTACTGGCTCCCTTTGTCCAGGACCCACCTGATCAAGAGCCTTAAGAGAGAAACC 360
Db 301 GCTT-----GAACCCACTGATGAAGAGCCTTAAGAGAGAAACC 339
Qy 361 ACCCACTAAAGTCGGAATCTTACACCTGATCAGAAAGAGAGATGATCAGGGTGCAGC 420
Db 340 ACCCACTAAAGTCGGAATCTTACACCTGATCAGAAAGAGAGATGATCAGGGTGCAGC 399
Qy 421 TGAGATTCAAGTGCCTGAGCCGATCCAGGAGTATGTCAGACAAAGACTGG 480
Db 400 TGAGATTCAAGTGCCTGAGCCGATCCAGGAGTATGTCAGACAAAGACTGG 459
Qy 481 CGATGATGTGAAGTGGTACTGATGTCAGAGGGGAAGATTCTACCAAAGCAGAGCACTT 540
Db 460 CGATGATGTGAAGTGGTACTGATGTCAGAGGGGAAGATTCTACCAAAGCAGAGCACTT 519
Qy 541 TAAATGCCAAGACGAGTGAAGGGAATCACAGCTTTAAGGAAGATAAGCTGAACAA 600
Db 520 TAAATGCCAAGACGAGTGAAGGGAATCACAGCTTTAAGGAAGATAAGCTGAACAA 579
Qy 601 C 601
Db 580 C 580

RESULT 5
AAD14983
ID AAD14983 standard; DNA; 611 BP.
XX
AC AAD14983;
XX
DT 01-NOV-2001 (first entry)
DE
DE Human NOV4 DNA.
XX
KW Human; NOVX; G-antigen; GAGE-like protein; interferon;
KW G-protein coupled receptor; GPCR; hepatocyte nuclear factor;
KW mast cell protease; gene therapy; proliferative disorder; cancer;
KW immune disorder; hepatic disorder; cirrhosis; viral infection;
KW hepatitis; neuroolfactory system-related disorder; neurological disorder;
KW Parkinson's disease; infertility; autoimmune disease; arthritis;
KW multiple sclerosis; allergy; wound healing; cytostatic; nootropic;
KW immunosuppressive; neuroprotective; vulnerary; hepatotropic; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..173
FT FT /*tag= a
FT CDS 174..521
FT FT /*tag= b
FT FT /product= "Human NOV4 protein"
FT 3'UTR 522..611
FT FT /*tag= c
XX
XX WO200161009-A2.
XX
XX 23-AUG-2001.
XX
XX 15-FEB-2001; 2001WO-US04828.
XX
XX 15-FEB-2000; 2000US-0182723.
XX 15-FEB-2000; 2000US-0182724.
XX 15-FEB-2000; 2000US-0182733.
XX 22-FEB-2000; 2000US-0183896.
XX 23-FEB-2000; 2000US-0184275.

```

```

PR 23-FEB-2000; 2000US-0184482.
PR 23-FEB-2000; 2000US-0184497.
PR 24-FEB-2000; 2000US-0184744.
PR 13-APR-2000; 2000US-0197083.
PR 10-AUG-2000; 2000US-0224157.
PR 18-SEP-2000; 2000US-0233405.
PR 27-SEP-2000; 2000US-0236060.
PR 02-JAN-2001; 2001US-0259414.
PR 18-JAN-2001; 2001US-0262454.
PR 14-FEB-2001; 2001US-0783429.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Malyankar UM, Tchernev VT, Padigar M, Taupier RJ, Spytek KA;
XX Majumder K, Guo X, Spaderna SK, Boldog FL;
PI
PI
XX
XX WPI; 2001-514775/56.
XX P-PSDB; AAE08583.
XX
XX Isolated novel polypeptides useful for diagnosis of and treating
XX cancer, infertility, autoimmune diseases, arthritis, multiple
XX sclerosis, allergies, wound healing and hepatic disorders -
XX
XX Claim 9; Page 14; 140pp; English.
XX
XX The present sequence is a human NOV4 DNA. The NOVX protein has homology
XX with one of G-antigen (GAGE)-like protein, interferon, G-protein coupled
XX receptor (GPCR), hepatocyte nuclear factor or mast cell protease. The
XX NOVX is useful for treating or preventing a pathology associated with
XX NOVX. It is also useful for determining the presence or amount of NOVX
XX DNA in a sample, for identifying a potential therapeutic agent and in
XX gene therapy. It is also useful for determining the presence of or
XX predisposition to a disease associated with altered levels of NOVX. It is
XX also useful for the diagnosis and treatment of proliferative disorders,
XX e.g., cancer, immune disorders, hepatic disorders, e.g., cirrhosis, viral
XX infections, e.g., hepatitis, neuroolfactory system-related disorders,
XX neurological disorders, e.g., Parkinson's disease, infertility,
XX autoimmune diseases, arthritis, multiple sclerosis, allergies and wound
XX healing.
XX
XX Sequence 611 BP; 196 A; 127 C; 152 G; 135 T; 1 other;
SQ
Query Match 75.9%; Score 507.8; DB 22; Length 611;
Best Local Similarity 94.6%; Pred. No. 7.3e-144;
Matches 566; Conservative 0; Mismatches 8; Indels 24; Gaps 3;
Qy 60 GTTCACGTGGGCACTTCCCTTCGACCCCTTGGCCACGTGGTACCGCT-GGGAGCTGT 118
Db 9 GTACACTGGGCATCTTCCCTTCGACCCCTTGGCCACGTGGTACCGCTGGGGAGCTGT 68
Qy 119 GAGAGTGTGAGGGGACGTTCCAGCCGCTCTGGACTCTTTCTCTCTACTGAGACGAGCC 178
Db 69 GAGAGTGTGAGGGACAGCTTCCAGCCGCTCTGGACTCTTTCTCTACTGAGACGAGCC 128
Qy 179 TATAGTCCGAGCCAGTCTCTCCAGGAACCTGAATAGTGAATATGATTCGCGAGGA 238
Db 129 TATAGTCCGAGCCAGTCTCTCCAGGAACCTGAATAGTGAATATGATTCGCGAGGA 188
Qy 239 AGATCAACATATAGGCTTAGCCCTAGCCAAAGAAAGTTTACAGCCTCTCAGCTGATTTGGGGCT 298
Db 189 AGATCAACATATAGGCTTAGCCCTAGCCAAAGAAAGTTTACAGCCTCTCAGCTGATTTGGGGCT 248
Qy 299 ATGCTTACTGGCTCCCTTTGTCCAGGAACCCACTGATGAAGAGCCTTAAAGAGAGAA 358
Db 249 ATGCTT-----GAACCCACTGATGAAGAGCCTTAAAGAGAGAA 287
Qy 359 CCACCCACTAAAGTCGGAATCTTACACCTGA--TCAGAGAGAGAGATGATCAGGGTG 416
Db 288 CCACCCACTAAAGTCGGAATCTTACACCTGATCTNAGAGAGAGAGATGATCAGGGTG 347
Qy 417 CAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTGAGACAAAGA 476
Db 348 CAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTGAGACAAAGA 407

```

QY 477 CTGGGATGATGCTGAAGGTGGTACTGATGTCAGGGGGAAGATTCTTACCAAAAGCAGAGC 536
 Db 408 CTGGGATGATGCTGAAGGTGGTACTGATGTCAGGGGGAAGATTCTTACCAAAAGCAGAGC 467
 QY 537 ACTTTAAATGCCAGAGCAGGTGAAGGGAATACAGAGTTTAAAGGAAGATAGCTGAA 596
 Db 468 ACTTTAAATGCCAGAGCAGGTGAAGGGAATACAGAGTTTAAAGGAAGATAGCTGAA 527
 QY 597 ACAACAAACTCTTTTATATTAGATATTTTAAAGAGTCTTTAATAAATTTT 654
 Db 528 ACAACAAACTCTTTTATATTAGATATTTTAAAGAGTCTTTAATAAATTTT 585

RESULT 6

AAK52902/c
 ID AAK52902 standard; cDNA; 665 BP.

XX AC AAK52902;

XX DT 06-NOV-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 2431.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.

XX OS Homo sapiens.

XX PN WO200157190-A2.

XX PD 09-AUG-2001.

XX XX 05-FEB-2001; 2001WO-US04098.

XX XX 03-FEB-2000; 2000US-0496914.

XX XX 27-APR-2000; 2000US-0560875.

XX XX 20-JUN-2000; 2000US-0598075.

XX XX 19-JUL-2000; 2000US-0620325.

XX XX 01-SEP-2000; 2000US-0654936.

XX XX 15-SEP-2000; 2000US-0663561.

XX XX 20-OCT-2000; 2000US-0693325.

XX XX 30-NOV-2000; 2000US-0728422.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 FI Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZH;
 FI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX WPI; 2001-476283/51.

XX DR P-PSDB; AAM79769.

XX PT Nucleic acids encoding polypeptides with cytokine-like activities,
 FI useful in diagnosis and gene therapy -

XX PS Claim 1; Page 4711; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing

CC were missing at the time of publication.

XX SQ Sequence 665 BP; 156 A; 170 C; 139 G; 200 T; 0 other;

Query Match 70.1%; Score 469.2; DB 22; Length 665;
 Best Local Similarity 88.1%; Pred. No. 4.1e-132;
 Matches 580; Conservative 0; Mismatches 43; Indels 35; Gaps 5;

QY 11 GCACCGACTTCAGTGTGCATGTTCTCTTGGACACCTGCTCAGTGTGCTATTTCACTGGGC 70

Db 665 GCACCGACTTCAGTGTGCATGTTCTCTTGGACACCTGCTCAGTGTGCTATTTCACTGGGC 606

QY 71 ATCTTCTCCCTTCGACCCCTTTGCCACGCTGTGACCGCTGGGAGCTGTGAGAGTGTGAGG 130

Db 605 ATCTTCTCCCTTCGACCCCTTTGCCACGCTGTGACCGCTGGGAGCTGTGAGAGTGTGAGG 546

QY 131 GGCAGTTCCAGCCGTTCTGACTCTTTCTCTCTACTGACGACGAGCTTATAGTTCGCA 190

Db 545 GGCAGTTCCAGCCGTTCTGACTCTTTCTCTCTACTGACGACGAGCTTATAGTTCGCA 486

QY 191 GGCAGTCTCCAGGAACTGAAATAGTGAATATAGTTGGGAGGAGATCAACATAT 250

Db 485 GGCAGTCTCCAGGAACTGAAATAGTGAATATAGTTGGGAGGAGATCAACATAT 426

QY 251 AGGCTTAGGCCCAAGAAAGTTTACAGCTCTCTGAGCTGATGGGGCTATGCTTACGGC 310

Db 425 AGGCTTAGGCCCAAGAAAGTTTACAGCTCTCTGAGCTGATGGGGCTATGCTTACGGC 372

QY 311 TCCCTTTGTCCAGGAAACCCACTGATGAAGAGCCTTAAAGAGAAACCCACCTAAA 370

Db 371 TCCCTTTGTCCAGGAAACCCACTGATGAAGAGCCTTAAAGAGAAACCCACCTAAA 327

QY 371 AGTGGAAATCTCACCTGATCAGAGAGAGAGATGATCAGGG- - - - -TGCAGCTGAGA 425

Db 326 AGTGGAAATCTCACCTGATCAGAGAGAGAGATGATCAGGGTTGCAGCTTGAGATT 267

QY 426 TTCAGTGCCTGA- - - - -CCTGGAGCCGATCTCCAGGAGCTATGT- - - - -CAGACAAAGCTG 479

Db 266 TCAAGTGGCTTGAACCTGGGAGAGCCGATTTTCCAGGAGAGCTGTTGTCAGACCAAGACTG 207

QY 480 GGGATGGATGTG- - - - -AAGTGGTACTGATGTCAAGGGGAAAGATTCTACCAAAAGCAGAGC 536

Db 206 GGGATGGTGTGTAAGTGGTACTGATGTCAAGGGGAAAGATTCTACCAAAAGCAGAGC 147

QY 537 ACTTTAAATGCCAGAGCAGGTGAAGGAAATACAGAGTTTAAAGGAAGATAGCTGAA 596

Db 146 ACTTTAAATGCCAGAGCAGGTGAAGGAAATACAGAGTTTAAAGGAAGATAGCTGAA 87

QY 597 ACAACACAACTGTTTTTATATTAGATATTTTAAAGAGCTTTAATAAATTTT 654

Db 86 ACAACACAACTGTTTTTATATTAGATATTTTAAAGAGCTTTAATAAATTTT 29

RESULT 7

AAI58744

ID AAI58744 standard; cDNA; 659 BP.

XX AC AAI58744;

XX DT 22-OCT-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 947.

XX Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
XX P-PSDB; AAM39588.
XX PT Novel nucleic acids and polypeptides, useful for treating disorders
XX PT such as central nervous system injuries -
XX PS Claim 1; SEQ ID NO 947; 10078pp; English.
XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX CC the encoded polypeptides (AAM38642-AAM42213) with nootropic
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide
XX CC of the invention may be used to treat diseases of the peripheral nervous
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and
XX CC localized neuropathies and central nervous system diseases, such as
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX CC utilisation of the activities such as: immune system suppression,
XX CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX CC assays for receptor activity, arthritis and inflammation, leukaemias and
XX CC C.N.S disorders.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification.
XX SQ Sequence 659 BP; 197 A; 137 C; 173 G; 152 T; 0 other;
Query Match 47.4%; Score 317; DB 22; Length 659;
Best Local Similarity 77.4%; Pred. No. 7.6e-86;
Matches 442; Conservative 0; Mismatches 85; Indels 44; Gaps 3;
QY 86 CTTTGGCCACGTGTGACCGCTGGGAGCTGTGAGAGTGTGAGGGGACGCTCCAGCG 145
DB 109 CTTTTCACCGTGGTGAATGCCCTGGAGTGTGAGGGTGTGAGGTGCGCTTCTGCTG 168
QY 146 TCTGACTCTTTCTCTCTACTGAGACGCGAGCTATAGTCCGACGGCCAGTCTCCAG 205
DB 169 TCTGACTTTTCTGCTCCACTGAGACGCGAGCTGATTCTGT----- 210
QY 206 GACTGAATAGTGAATATAGTGTGGCGAGAGATCAATATAGAGCTAGGCAAGA 265
DB 211 ----TTGCAGTGAATATGATTGGCGAGGAGATCAATATAGAGCTAGGCGAGG 266
QY 266 AGAAGTTACAGCTCTCTGAGCTGATTGGGGCTATGCTTACTGGCTCCCTTTGTCAG 325
DB 267 AGAAGTGTACCACTCTCTGAGCTGATTGGGCTATGCT-----G 305
QY 326 GAACCACTGTAGAGAGCTTAAGAGAGAAACCAACCCACTAAAGTGGGAATCTACA 385
DB 306 GAGCCCGGTGTAGGAGAGCTCAGCAAGAGGAACCAACCACTGAAGTGGGATCTGCA 365
QY 386 CTTGATCAGAAGAGAGATCATCAGGTGCGAGCTGAGATTCAAGTGCCTGACCTGAA 445

DB 366 CCGTGTCCAGGAGAGAGAGATCAGGCTGAGCTGAGACTCAAGTGCCTGACCTGGAA 425
QY 446 GCCGATCTCCAGAGCTATGTGACAAAGACTGGGATGGATGTGAAGGTGTACTGAT 505
DB 426 GCTGATCTCCAGAGCTGTCTCAGTCAAAGACTGGGGTGAATGTGGAATGTCTCTGAT 485
QY 506 GTCAAGGGGAAGATTCTACAAAAGCAGAGCACTTTTAAATGCCAGAAGCAGGTGAAGG 565
DB 486 GACACGGGAGAGATTCTGCCAAATCAGAACATTTAAATGCCAGAAGAGGTGACAGG 545
QY 566 AATACAGGTTTAAAGGAAGATAAGCTGAACCAACAC- AAAGTGTGTATATATAGATA 624
DB 546 CAACCCACAGGTTTAAATGAAGCAAGCTGAACCAACACCAAAAGCTGTTTTTATCTAAGATA 605
QY 625 TTTTACTTTAAAGAGCTCTTAATAAATTTTGG 655
DB 606 TTGACTTTAAATAATCGAAATAAATTTTGG 636
RESULT 8
AAI60530
ID AAI60530 standard; cDNA; 532 BP.
XX AC AAI60530;
XX DT 22-OCT-2001 (first entry)
XX DE Human polynucleotide SEQ ID NO 4519.
XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX KW peripheral nervous system; neuropathy; central nervous system; CNS;
XX KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX KW leukaemia; ss.
XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
XX P-PSDB; AAM41374.
XX PT Novel nucleic acids and polypeptides, useful for treating disorders
XX PT such as central nervous system injuries -
XX PS Claim 1; SEQ ID NO 4519; 10078pp; English.
XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide
XX CC of the invention may be used to treat diseases of the peripheral nervous
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and
XX CC localized neuropathies and central nervous system diseases, such as
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX CC utilisation of the activities such as: immune system suppression,
XX CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX CC assays for receptor activity, arthritis and inflammation, leukaemias and
XX CC C.N.S disorders.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification.

CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 532 BP; 154 A; 108 C; 150 G; 120 T; 0 other;

Query Match 44.3%; Score 296.2; DB 22; Length 532;
Best Local Similarity 76.9%; Pred. No. 1.4e-79;
Matches 439; Conservative 0; Mismatches 73; Indels 59; Gaps 4;
QY 87 CTTTGGCCAGCTGTGACCGCTGGGAGCTGTGAGAGTGTGAGGGCAGCTTCCAGCGT 146
Db 1 CTTTTCGCCAGTGTGAATGCCCTCGAGTGTGAGGGTGTGAGGTCGCGTTCTCTGCTGT 60
QY 147 CTGAGCTCTTCTCTCTACTAGAGCGAGCCTATAGGTCCGAGGCCAGTCTCCCCAGG 206
Db 61 CTGAGCTTCTTCT 94
QY 207 AACTGAAATAGTGAATATGAGTTGGCGAGGAAGATCAACATATAGGCTTAGGCCAAGAA 266
Db 95 -----GTGAAATATGATTTGGCGAGGAAGATCAACATATAGGCTTAGGCCGAGGA 144
QY 267 GAAGTTTACAGCTCTCTGAGCTGATTTGGGCTATGCTTACTGCTCCCTTTGTCCCAGG 326
Db 145 GAAGTGTACACCTCTGAGCTGATTTGGGCTATGCT-----GG 183
QY 327 AACCCACTGATGAAGAGCCTTAAGAGAGAACCCACCTAAAGTCGGAATCTCTACAC 386
Db 184 AGCCCGGTGATGAGGAGCCTCAGCAAGAGAGAACCCACCTAAAGTCGGAATCTCTGAC 243
QY 387 CTGATCAGAGAGAGAGATGATCAGGTGCAGCTGAGATTCAGTGCCTGACCTGGAG 446
Db 244 CTGATCAGAGAGAGAGAGATGATCAGGTGCAGCTGAGATTCAGTGCCTGACCTGGAG 303
QY 447 CCGATCTCCAGGAGCTATGTGACAGCAAGAGCTGGGATGATGTGAAGTGTGATGATG 506
Db 304 CTGATCTCCAGGAGCTCTCTCAGTCAAGAGCTGGGATGATGTGCGAGATGCTCTGATG 363
QY 507 TCAGGGGAGATTTCTACCAAGAGCAGAGCATTAAATGCCAGAA-CCAGGTGAGGG 565
Db 364 TCAGGGGAGATTTCTGACAAAGTCTGAGCAATTTAAATGCCAGAGGAGGTGACAGG 423
QY 566 AAATCAGAGTTTAAAGGAGAGATAAGCTGAAACAAACAC-AAACTGTTTTTATATAGATA 624
Db 424 CAACACAGGTTTAAATGAGAGAGAGCTGAAACAAACAACTGTTTTTATCTAAGATA 483
QY 625 TTTTACTTTTAAAGAGTCTTAAATTTT 655
Db 484 TTTGACTTAAATAATATCGAATAAATTTT 514

RESULT 9
ID ABX77605
XX ABX77605 standard; cDNA; 509 BP.
AC ABX77605;
XX
XX
DT 09-APR-2003 (first entry)
XX
DE Differentially expressed breast cancer associated cDNA #100.
XX
KW Breast cancer; differential gene expression; BC-cDNA;
KW breast cancer diagnosis; breast cancer monitoring;
KW breast cancer treatment; breast cancer staging; gene; ss.
XX
OS Unidentified.

XX US2002156263-A1.
XX 24-OCT-2002.
XX 04-OCT-2001; 2001US-0974298.
XX 05-OCT-2000; 2000US-238331P.
XX (CHEN/) CHEN H.
XX Chen H;
XX WPI; 2003-182653/18.
XX New cDNAs, which are differentially expressed in (metastatic) breast
XX cancer useful for diagnosing or staging, breast cancer, or for
XX monitoring the treatment of breast cancer in an individual
XX
XX Claim 1; SEQ ID NO 126; 30pp; English.
XX
XX The invention describes a combination of cDNAs (designated BC-cDNAs),
XX which are differentially expressed in breast cancer. The combination
XX includes 152 cDNA sequences, or their complements. The protein encoded
XX by any of these BC-cDNAs is useful for screening several molecules or
XX compounds to identify at least one ligand that specifically binds the
XX protein, producing or preparing polyclonal or monoclonal antibodies, or
XX purifying antibodies from a sample. The antibodies, which specifically
XX bind the protein differentially expressed in breast cancer is useful for
XX detecting the expression of a protein in a sample. The BC-cDNAs are
XX also useful for diagnosing, monitoring the treatment of, or staging,
XX breast cancer. This sequence represents a differentially expressed
XX breast cancer associated cDNA.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from the US patent office at
XX seqdata.uspto.gov/sequence.html?docID=20020156263.
XX
XX Sequence 509 BP; 156 A; 100 C; 139 G; 112 T; 0 other;

Query Match 43.4%; Score 290.2; DB 25; Length 509;
Best Local Similarity 76.9%; Pred. No. 9.3e-78;
Matches 419; Conservative 0; Mismatches 68; Indels 58; Gaps 3;
QY 112 GAGCTGTGAGAGTGTGAGGGGACGTTCCAGCCGCTGAGACTCTTTCTCTCTACTGAGA 171
Db 1 GAGTTGTGAGGGTGTGAGGGTCCGTTCTCTGCTCTGAGACTTTTCTGCTCCACTGAGA 60
QY 172 CGCAGCCTATAGGTCCGAGGCCAGTCTCCAGGAACCTGAAATAGTGAATATGAGTTG 231
Db 61 CGCAGCTCT-----GTGAAATATGATTTG 84
QY 232 CGCAGGAGATCAACATATAGGCTTAGGCCAAGAGAGTTTACAGCTCTCTGAGCTGAT 291
Db 85 CGCAGGAGATCAACATATAGGCTTAGGCCAAGAGAGTTGTTACCACTCTCTGAGCTGAT 144
QY 292 TGGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCACTGATGAAGAGCCTAAAGA 351
Db 145 TGGGGCTATGCT-----GGAGCCCGGTGATGAGGAGCCTCAGCA 183
QY 352 AGAGGAACCCACTAAAGTCCGAAATCTTACACCTGATCAGAGAGAGAGATGATCA 411
Db 184 AGAGGAACCCACTAAAGTCCGAAATCTTACACCTGATCAGAGAGAGAGAGATGATCA 243
QY 412 GGGTGCAGCTGAGATTTCAAGTCCCTGACCTGGAAAGCCGATCTCCAGGAGCTATCTCAGAC 471
Db 244 GGGTGCAGCTGAGATTTCAAGTCCCTGACCTGGAAAGCTGATCTCCAGGAGCTGTCTCAGTC 303
QY 472 AAAGACTGGGATGATGTGAAGTGTGATGTCTCAAGGGGAGAGATTTCTACCAAGGC 531
Db 304 AAAGACTGGGATGATGTGAAGTGTGATGTCTCAAGGGGAGAGATTTCTCCCAAAATC 363
QY 532 AGAGCACTTTAAATGCCAGAGAGAGGAGGAAATACACAGGTTTAAAGGAAGATAAG 591

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (II) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (III). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 503 BP; 158 A; 98 C; 121 G; 105 T; 21 other;

Query Match 42.0%; Score 280.8; DB 23; Length 503;
Best Local Similarity 81.1%; Pred. No. 6.7e-75;
Matches 360; Conservative 0; Mismatches 62; Indels 22; Gaps 2;

QY 213 AATAGTGAATATGAGTGGCGAGGAAGATCAACATATAGGCTAGGCCAAGAAGATT 272
DB 49 AATGTGAATATGATTTGGCGAGGAAGATCAACATATAGGCTAGGCCGAGAGAAGTG 108

QY 273 TACAGCTCTGAGCTGATGGGCTATGCTTACTGGCTCCCTTTGTCGCCAGAACCA 332
DB 109 TACCACCTCTGAGCTGATGGGCTATGCT-----GGAGCCCG 147

QY 333 CTGATCAAGAGCTTAAGAGAGAGAAACCCACCTAAAGTGGGAATCTACACCTGATC 392
DB 148 GTGATGAGAGCTCTGACAGAGAGAACCCAACTGAAAGTGGGATCTGCACCTGGTC 207

QY 393 AGAAGAGAGAGATGATGAGGTGCGAGCTGAGATTCAAGTGCCTGACCTGGAGCGATC 452
DB 208 AGGAGAGAAAGAGATCAGGGTTGAGCTCAAGTCAAGTGCCTGACCTGGAGCTGATC 267

QY 453 TCCAGAGGCTATCTCAGACAAAGACTGGGATGGATGATGAGTGTGATGATGATCAAG 512
DB 268 TCCAGGAGCTGTCTCAGTCAAGACTGGGGTGAATGTGGAATGTCTGATGACGAG 327

QY 513 GGAAGATTCTACCAAGAGAGACACTTTTAAATGCCAGAGCAGGTGAAGGGAATCAC 572
DB 328 GGAAGATTCTGCCAAATATCAGAACAAATTTAAATGCCAGAGGAGGTGACAGGCAACC 387

QY 573 AGTTTAAAGGAAGATGAGCTGAAACACAC-AAAACCTGTTTTTATATTAGATATTTACT 631
DB 388 AGTTTAAATGAGAGAGCTGAAACACCAACCAACCAACCTGTTTTTATTAAGATATTTGACT 447

QY 632 TTAAGAGCTCTTAATAAATTTTG 655
DB 448 TAAATATATCAATAAATTTTG 471

RESULT 12
AAS69486
ID AAS69486 standard; cDNA; 661 BP.
XX
AC AAS69486;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #5290.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.
XX OS
XX PN WO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WO-US08631.
XX
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PI Drmanac RT, Liu C, Tang YT;
XX P-PSDB; ABG05299.
XX
XX DR WPI; 2001-639362/73.
XX
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity
XX
XX PS Claim 1; SEQ ID No 5290; 103pp; English.
XX
XX CC The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (III). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 661 BP; 212 A; 129 C; 158 G; 162 T; 0 other;

Query Match 36.2%; Score 242.4; DB 23; Length 661;
Best Local Similarity 81.5%; Pred. No. 3.6e-63;
Matches 362; Conservative 0; Mismatches 56; Indels 26; Gaps 5;

QY 217 GTGAATATGATGTTGGCGAGGAAGATCAACATATAGGCTAGGCCAAGAAGATTACA 276
DB 222 GTGAATATGATGTTGGCGAGGAAGATCAACATATAGGCTAGGCCAAGAAGATTACC 281

QY 277 GCCTCTCAGCTGATTTGGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCACTGA 336
DB 282 ACCTCTCAGCTGATTTGGGGCTATGCT-----GGAGCCCGGTGA 320

QY 337 TGAAGAGCTTAAGAGAGGAAGAACCCACCTAAAGTGGGAATCCTACACCT-GATCAGA 395
DB 321 TGAGAGCTCTCAGCAAGAGAGGAAGAACCCACCTAAAGTGGGAATCCTGACCTGGGTGAG 380

QY 396 AGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGTGCCTGACCTGAGAA-GCCGATCTC 454
DB 381 AGAGAGAGAGAGATCAGGGTGCAGCTGAGATTCAGTGCCTGACCTGAGAGGCTGATCTC 440

QY 455 CAGGAGCTATGTACAGCAAGAGCTGGGATGGATGAGTGGTGTACTGATGTCAAGGGG 514
DB 441 CAGGAGCTGTCTCAGTCAAGAGACTGGGGGTGAATGTGGAAATGTTCTGTATGACACGGG 500

QY 515 AAGATTCTACCAAGCAGAGC-ACCTTAAATGCCAGAA-GCAGGTGAGGGGAATCAC 572
D501 AAGATTCTGCCAAATCAGAACATTTTAAATGCCAGAGGGAGGTGACAGGCAACCAC 560
QY 573 AGGTTTAAAGGAGATAGCTGAAACACAC-AAACTGTTTTTATATTAGATATTTTACT 631
D561 AGGTTTAAATGAAGCAGCTGAAACACACAAACAACTGTTTTTATCTAAGATATTTGACT 620
QY 632 TTAAGAGCTCTTAATAATTTTTG 655
D621 TAAAAATATCGAATAAACTTTTG 644

RESULT 13
AAH83260
ID AAH83260 standard; cDNA; 320 BP.
XX
AC AAH83260;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human ovarian tumour associated polynucleotide sequence SEQ ID NO:884.
XX
KW Human; ovarian tumour; ovarian cancer; diagnosis; gene therapy;
KW immunogenic; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200151513-A2.
XX
PD 19-JUL-2001.
XX
PF 16-JAN-2001; 2001WO-US01575.
XX
PR 14-JAN-2000; 2000US-0176722.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA;
XX
DR WPI; 2001-425866/45.
XX

Novel ovarian tumor proteins, and nucleic acids encoding them, used to treat and diagnose cancers, particularly ovarian cancer -
Claim 5; Page 225; 338pp; English.

AAH82377 to AAH83878 represent human ovarian tumour-associated polynucleotide sequences which encode ovarian tumour proteins. The ovarian tumour protein and polynucleotide sequences have cytostatic activity, and can be used in gene therapy and vaccine production. The ovarian tumour proteins and polynucleotides can be used to inhibit the development of cancer, particularly ovarian cancer. They can also be used to diagnose the onset and progression of cancer.

Sequence 320 BP; 96 A; 65 C; 85 G; 74 T; 0 other;
Query Match 32.1%; Score 215; DB 22; Length 320;
Best Local Similarity 100.0%; Pred. No. 5.3e-55;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 420 CTGAGATTCAAGTGGCTGACCTGGAGCGGATCTCCAGGAGCTATGTCCAGCAAAAGACTG 479
D1 CTGAGATTCAAGTGGCTGACCTGGAGCGGATCTCCAGGAGCTATGTCCAGCAAAAGACTG 60
QY 480 GGGATGGATGTAAGGTGCTATGATGTCAGGGGAAGATCTACCAAAAGCAGGCACT 539
D61 GGGATGGATGTAAGGTGCTATGATGTCAGGGGAAGATCTACCAAAAGCAGGCACT 120
QY 540 TTAATAATGCCAAGCAGGTGAAGGGAATCAGAGTTTAAAGGGAAGATAAGCTGAAACA 599
D121 TTAATAATGCCAAGCAGGTGAAGGGAATCAGAGTTTAAAGGGAAGATAAGCTGAAACA 180

QY 600 ACACAACTGTTTTTATATTAGATATTTTACTTTA 634
D181 ACACAACTGTTTTTATATTAGATATTTTACTTTA 215

RESULT 14
AAS24637
ID AAS24637 standard; cDNA; 214 BP.
XX
AC AAS24637;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human ovarian PCR-subtracted cDNA library clone #818.
XX
KW Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;
KW gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;
KW primer; probe.
XX
OS Homo sapiens.
XX
PN WO200157207-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03733.
XX
PR 04-FEB-2000; 2000US-0180403.
XX
PR 28-MAR-2000; 2000US-0192745.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Mannion J;
XX
DR WPI; 2001-488879/53.
XX

New polynucleotides encoding ovarian tumour proteins, useful for treating ovarian cancer, and as probes, primers, and markers of cancer progression -

Example 1; page 241-242; 378pp; English.

The invention comprises compositions used for the therapy and diagnosis of ovarian cancer. The compositions comprise one or more ovarian tumour proteins, their associated polynucleotides, or immunogenic portions of the proteins. The ovarian tumour polynucleotides and polypeptides are useful for stimulating and/or expanding T cells specific for a tumour protein. They are also useful for inhibiting the development of cancer in a patient with an ovarian tumour DNA or protein by incubating isolated T-cells allowing them to proliferate, and administering to the patient. The sequences can be used as markers for cancer, for example, to monitor ovarian cancer progression. Probes and primers are useful in nucleic acid hybridisation, in detecting the presence of complementary sequences in a given sample, for preparing mutant species and for preparing other genetic constructions. Sequences AAS23820-AAS25231 and AAS25238-AAS25549 represent human ovarian tumour protein cDNA clones.

Sequence 214 BP; 74 A; 34 C; 55 G; 51 T; 0 other;

Query Match 32.0%; Score 214; DB 22; Length 214;
Best Local Similarity 100.0%; Pred. No. 8.7e-55;
Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 420 CTGAGATTCAAGTGGCTGACCTGGAGCGGATCTCCAGGAGCTATGTCCAGCAAAAGACTG 479
D1 CTGAGATTCAAGTGGCTGACCTGGAGCGGATCTCCAGGAGCTATGTCCAGCAAAAGACTG 60
QY 480 GGGATGGATGTAAGGTGCTATGATGTCAGGGGAAGATCTACCAAAAGCAGGCACT 539
D61 GGGATGGATGTAAGGTGCTATGATGTCAGGGGAAGATCTACCAAAAGCAGGCACT 120
QY 540 TTAATAATGCCAAGCAGGTGAAGGGAATCAGAGTTTAAAGGGAAGATAAGCTGAAACA 599

Db 121 TTAATGCCAGAGCAGGTGAAGGAAATACACAGGTTTAAAGGAAGATAAGCTGAACA 180
Qy 600 ACACAACTGTTTATATAGATATTTACTTT 633
Db 181 ACACAACTGTTTATATAGATATTTACTTT 214

RESULT 15
AAF68151
ID AAF68151 standard; cDNA; 399 BP.
XX
AC AAF68151;
XX
DT 12-APR-2001 (first entry)
XX
DE Human lung tumour protein related nucleotide sequence SEQ ID NO:69.
XX
KW Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
KW lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;
KW cytostatic; antisense inhibition; ss.
XX
OS Homo sapiens.
XX
XX WO200100828-A2.
PN
XX 04-JAN-2001.
XX
XX 30-JUN-2000; 2000WO-US18061.
XX
XX 30-JUN-1999; 99US-0346492.
PR 15-OCT-1999; 99US-0419356.
PR 17-DEC-1999; 99US-0466867.
PR 30-DEC-1999; 99US-0476300.
PR 06-MAR-2000; 2000US-0519642.
PR 22-MAR-2000; 2000US-0533077.
PR 10-APR-2000; 2000US-0546259.
PR 27-APR-2000; 2000US-0560406.
PR 05-JUN-2000; 2000US-0589184.
XX
XX (CORI-) CORIXA CORP.
XX
XX Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
PI Retter MW, Mannion J;
XX
XX WPI; 2001-071486/08.
XX
XX Lung tumor-associated proteins and the nucleic acids that encode them,
XX useful for preventing, diagnosing and treating lung cancer -
XX
XX Claim 4; Page 173; 436pp; English.

The present invention describes immunogenic portions of lung tumour-associated proteins (I) and the nucleic acids (NAs) that encode them. (I) have cytostatic activity and can be used in gene therapy, antisense inhibition and in vaccines. The NAs and the lung tumour-associated proteins they encode may be used in the prevention, treatment and diagnosis of diseases associated with their inappropriate expression, especially lung cancers. For example, the NAs may be administered to treat diseases by rectifying mutations or deletions in a patient's genome that affect the activity of the protein by expressing inactive proteins or to supplement the patient's own production of (I). Additionally, the NAs may be used to produce the lung-tumour associated protein, according to standard recombinant DNA methodology. Conversely, antisense NA molecules may be administered to down regulate protein expression by binding with the cells own genes and preventing their expression. The NA and complementary sequences may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar NA sequences in samples, and hence which patients may be in need of treatment for lung cancer. The (I) may be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of the expression and activity of the protein. AAF68083 to AAF68878 and AAB76848 to AAB76878 represent human lung tumour protein related

CC nucleotide and protein sequences which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 399 BP; 150 A; 87 C; 94 G; 67 T; 1 other;
Query Match 30.5%; Score 204; DB 22; Length 399;
Best Local Similarity 77.4%; Pred. No. 1.3e-51;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
Qy 319 GTCCAGGAAACCCACTGATGAAGAGCTTAAAGAGAGAAACACCCACTAAAGTGGAA 378
Db 40 GTCCAGGAGCCAGTAATGAGAGGCCCAAAAGAAAGAACCCAGCAGCTGAAGTCGGGA 99
Qy 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGCTGAGCTGAGATTCAGTGCCTGA 438
Db 100 TCCTACACCTGGGAGAGAGAGATCAGGATCAGCTGAGATCCAGTGGCGGA 159
Qy 439 CCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGGATGGATGCAAGGT-G 497
Db 160 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACCCGGGATAAATCTGGATTGG 219
Qy 498 GTACTGATGTCGAAGGGGAAGATTCTACCAAAAGCAGAGCACTTTAAAATGCCAGAAAGCAG 557
Db 220 GTTCCGGCTCAAGGTGAAGATTAATACCTTAAAGAGAACACTGTAAATGCCAGAAAGCAG 279
Qy 558 GTGAAGGGAATCACAGGTTTAAAGGAAGATAAGCTGAAACACACAAACTGTTTTTATA 617
Db 280 GTGAAGAGCAACCAAGTTTAAATGAAGCAAGCTGAAACACCAAGCTGGTTTATA 339
Qy 618 TTAGATATTTTACTTTTAAAGAGCTTTAATAATTTT 653
Db 340 TTAGATATTTTACTTTTAAAGAGCTTTTCAATAAAGTTTT 375

RESULT 16
ABK38062
ID ABK38062 standard; cDNA; 399 BP.
XX
AC ABK38062;
XX
XX 21-MAY-2002 (first entry)
DT
DE cDNA encoding clone #19107 (L552S) of lung tumour protein.
XX
XX Lung tumour; cancer; T cell; immune response stimulator;
KW cytostatic; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200204514-A2.
XX
XX 17-JAN-2002.
XX
XX 10-JUL-2001; 2001WO-US22059.
XX
XX 11-JUL-2000; 2000US-0614124.
PR 29-AUG-2000; 2000US-0651563.
PR 08-SEP-2000; 2000US-0658824.
PR 26-SEP-2000; 2000US-0671325.
PR 06-OCT-2000; 2000US-0677419.
PR 30-OCT-2000; 2000US-0702705.
PR 13-DEC-2000; 2000US-0736457.
PR 03-MAY-2001; 2001US-0849626.
XX
XX (CORI-) CORIXA CORP.
XX
XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
PI Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS;
PI McNabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;
XX
XX WPI; 2002-164634/21.
XX
XX Novel polynucleotide encoding a lung tumour polypeptide useful for

```
PT stimulating and/or expanding T cells specific for a tumour protein -
XX
XX Example 1; SEQ ID No 69; 223pp; English.
XX
CC The invention describes an isolated polynucleotide and polypeptide
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein for determining the presence of a cancer in a patient. A
CC composition containing the polynucleotide and/or polypeptide is useful
CC for treating a lung cancer in a patient. The polypeptide is useful for
CC removing tumour cells from a biological sample. The polynucleotide is
CC also useful as probe or primer to detect the level of mRNA encoding a
CC tumour protein. This sequence encodes a lung tumour associated protein
CC or protein fragment, described in the method of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 399 BP; 150 A; 87 C; 94 G; 67 T; 1 other;
SQ
Query Match 30.5%; Score 204; DB 24; Length 399;
Best Local Similarity 77.4%; Pred. No. 1.3e-51;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGAACCCACTGATGAAGAGCCCTAAAGAGAGAAACCCACCTAAAGTCGGAA 378
Db 40 GTCCAGGAGCCAGTAATGGAGAGCCCAAAAGAGAACCCAGCAGCTGAAAGTCGGGA 99
QY 379 TCCTACCTGATCAGAGAGAGAGATGATCAGGGTCAGCTGAGATTCAGTCCCTGA 438
Db 100 TCCTACCTGCGGACGACAGAGAGATGATCAGATCAGCTGAGATCCAGTCGGGA 159
QY 439 CTGGAAGCCGATCTCCAGGAGCTATGTGACAGAAAGACTGGGGATGATGTAAGGT-G 497
Db 160 CATGGAAGGTGATCTGCAAGAGCTGATCAGTCAGTCAAAACCCGGGATAATCTGGATTGG 219
QY 498 GTACTGATGTCAGGGAGAGATCTTACCAAGACGAGACATTTAAATGCCAGAGCAG 557
Db 220 GTTCGGGCGTCAAGGTGAAGATAATACCTAAAGAGGAACACTGTAAATGCCAGAGCAG 279
QY 558 GTGAAGGGAATACACAGGTTTAAAGAGAGATAAGCTGAAACACAACTGTTTTTATA 617
Db 280 GTGAAGAGCAACCAACAGTTTAAATGATGAGACAGCTGAACACCGCAAGCTGGTTTATA 339
QY 618 TTAGATATTTTACTTTAAAGAGTCTTAAATATTTT 653
Db 340 TTAGATATTTGACTTAAACTATCTCAATAAGTTTT 375
RESULT 17
ID ACA10391 standard; cDNA; 399 BP.
XX
XX ACA10391;
AC
XX
XX 05-JUN-2003 (first entry)
DT
XX
XX Human lung cancer-associated cDNA, L552S.
DE
XX
XX Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
XX
XX T cell expansion; CD4; CD8.
KW
XX
XX Homo sapiens.
OS
XX
XX US2002197669-A1.
PN
XX
XX 26-DEC-2002.
PD
XX
XX 03-MAY-2001; 2001US-0849626.
PF
XX
XX 13-DEC-2000; 2000US-0736457.
PR
XX
XX (BANG/) BANGUR C S.
PA
```

```
PA (FANG/) FANGER G R.
PA (WANG/) WANG A.
PA (WANG/) WANG T.
PA (SWIT/) SWITZER A P.
PA (MCNE/) MCNEILL P D.
PA (CLAP/) CLAPPER J D.
XX
XX Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;
PI Clapper JD;
PI
XX WPI; 2003-352750/33.
DR
XX
XX Novel lung cancer polynucleotide encoding lung cancer protein, useful
XX for detecting the presence of lung cancer in a patient, and in
XX pharmaceutical compositions, e.g. vaccines, for treating lung cancer
XX
XX Example 1; Page -: 72pp; English.
XX
XX The invention relates to a polynucleotide encoding a lung tumour protein,
XX comprising a sequence selected from any of the 14 sequences
XX mentioned in the specification, or a sequence (S2) mentioned in
XX specification, complement of S1, sequences consisting of at least 20
XX contiguous residues of S1, sequences that hybridise to S1, sequences
XX having 75%, preferably 90%, identity to S1, or degenerate variants of
XX S1. Also included are an isolated polypeptide (comprising a sequence (S3)
XX selected from any one of the 4 amino acid sequences mentioned in the
XX specification, a sequence encoded by the polynucleotide, or sequences
XX having at least 70%, preferably 90%, identity to a sequence encoded by
XX the polynucleotide), an expression vector comprising the polynucleotide
XX operably linked to an expression control sequence, a host cell
XX transformed or transfected with the vector, an isolated antibody (or its
XX antigen-binding fragment) that specifically binds to the polypeptide,
XX detecting the presence of a cancer in a patient, a fusion protein
XX comprising the polypeptide, an oligonucleotide that hybridises to
XX S1 under moderately stringent conditions, stimulating and/or expanding T
XX cells specific for a tumour protein (comprising contacting T cells with
XX the polynucleotide, protein or antigen-presenting cells, under conditions
XX and for a time sufficient to permit the stimulation and/or expansion of T
XX cells) and inhibiting the development of a cancer in a patient (by
XX incubating CD4+ and/or CD8+ T cells isolated from a patient with the
XX polynucleotide, protein or antigen presenting cells that express the
XX polynucleotide, such that T cells proliferate, administering to the
XX patient an effective amount of the proliferated T cells, and thus
XX inhibiting the development of a cancer in the patient. The
XX polynucleotide, protein and cells are useful in a composition for
XX stimulating an immune response in a patient, and for treating a cancer in
XX a patient (particularly lung cancer). The oligonucleotide is useful for
XX determining the presence of a cancer in a patient. The protein and
XX oligonucleotides are useful in pharmaceutical compositions, e.g.
XX vaccines. The polynucleotide is also useful as a probe or primer for
XX nucleic acid hybridisation, and in the design and preparation of
XX ribozyme molecules for inhibiting expression of tumour polypeptides and
XX proteins in tumour cells. An amplified portion of the polynucleotide is
XX useful for isolating a full-length gene from a suitable library.
XX The present sequence is a cDNA (full length, extended or partial)
XX isolated from a library derived from lung tumour/cancer cells.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from the USPTO
XX at seqdata.uspto.gov/sequence.html?DocId=20020197669.
XX
XX Sequence 399 BP; 150 A; 87 C; 94 G; 67 T; 1 other;
SQ
Query Match 30.5%; Score 204; DB 25; Length 399;
Best Local Similarity 77.4%; Pred. No. 1.3e-51;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGAACCCACTGATGAAGAGCTTAAAGAGAGAAACCCACCTAAAGTCGGAA 378
Db 40 GTCCAGGAGCCAGTAATGGAGAGCCCAAAAGAGAACCCAGCAGCTGAAAGTCGGGA 99
QY 379 TCCTACCTGATCAGAGAGAGAGATGATCAGGGTCAGCTGAGATTCAGTCCCTGA 438
```


XX Lung tumor-associated proteins and the nucleic acids that encode them,
PT useful for preventing, diagnosing and treating lung cancer -
XX
XX
XX Claim 4; Page 412; 436pp; English.
XX
XX The present invention describes immunogenic portions of lung tumour-
CC associated proteins (I) and the nucleic acids (NAs) that encode them.
CC (I) have cytostatic activity and can be used in gene therapy, antisense
CC inhibition and in vaccines. The NAs and the lung tumour-associated
CC proteins they encode may be used in the prevention, treatment and
CC diagnosis of diseases associated with their inappropriate expression,
CC especially lung cancers. For example, the NAs may be administered to
CC treat diseases by rectifying mutations or deletions in a patient's genome
CC that affect the activity of the protein by expressing inactive proteins
CC or to supplement the patient's own production of (I). Additionally, the
CC NAs may be used to produce the lung-tumour associated protein, according
CC to standard recombinant DNA methodology. Conversely, antisense NA
CC molecules may be administered to down regulate protein expression by
CC binding with the cells own genes and preventing their expression. The NA
CC and complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar NA sequences in
CC samples, and hence which patients may be in need of treatment for lung
CC cancer. The (I) may be used as antigens in the production of antibodies
CC and in assays to identify modulators (agonists and antagonists) of the
CC expression and activity of the protein. AAF68083 to AAF68878 and
CC AAB76848 to AAB76878 represent human lung tumour protein related
CC nucleotide and protein sequences which are used in the exemplification
CC of the present invention.
XX
XX Sequence 457 BP; 146 A; 105 C; 121 G; 85 T; 0 other;
SQ

Query Match 30.5%; Score 204; DB 22; Length 457;
Best Local Similarity 77.4%; Pred. No. 1.4e-51;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
XX 319 GTCCAGGAAACCATGTAGAGAGCCCTTAAGAGAGAACCCACCTAAAGTCGGAA 378
DB 120 GTCCAGGAGCCCGAGTAATGGAGAGCCCAAAAGAGAGACGAGCTGAAAGTCGGGA 179
XX 379 TCTACACCTGTATCAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGTGCCTGA 438
DB 180 TCTACACCTGGCAGCAGACAGAGAGATCAGATACAGTCCAGTCCAGTGGCGGA 239
XX 439 CTGGAAGCCGATCTCCAGGAGCTATGTGAGCAAAAGACTGGGGATGGATGTGAAGGT-G 497
DB 240 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGG 299
XX 498 GTACTGATCTCAAGGGGAGATTTCTACCAAAAGCAGAGACTTTAAATGCCAGAGAGCAG 557
DB 300 GTTCCGGCTCAAGGTGAAGATAATACCTTAAGAGAGACATCTGTAATGCCAGAGCAG 359
XX 558 GTCAAGGGAAATCACAGGTTTAAAGGAAGATAAGCTGAAACACACAAACTGTTTTATA 617
DB 360 GTGAAGAGCAACCAAGTTTAAATGAAGACAAGCTGAAACCAACGCAAGCTGGTTTTATA 419
XX 618 TTAGATATTTTACTTTAAGAGCTCTTAATAAATTTT 653
DB 420 TTAGATATTTGACTTAAACTATCTCAATAAAGTTTT 455

RESULT 20
ABK38762
ID ABK38762 standard; cDNA; 457 BP.
XX
XX ABK38762;
AC
XX
XX 21-MAY-2002 (first entry)
DT
XX cDNA encoding L552S lung tumour protein.
DE
XX Lung tumour; cancer; T cell; immune response stimulator;
XX cytostatic; gene; ss.
XX

XX Homo sapiens.
XX WO200204514-A2.
XX
XX 17-JAN-2002.
XX
XX 10-JUL-2001; 2001WO-US22058.
XX
XX 11-JUL-2000; 2000US-0614124.
XX 29-AUG-2000; 2000US-0651563.
XX 08-SEP-2000; 2000US-0658824.
XX 26-SEP-2000; 2000US-0671325.
XX 05-OCT-2000; 2000US-0677419.
XX 30-OCT-2000; 2000US-0702705.
XX 13-DEC-2000; 2000US-0736457.
XX 03-MAY-2001; 2001US-0849626.
XX (CORI-) CORIXA CORP.
XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW,
XX Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS;
XX McNabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;
XX
XX WPI; 2002-164634/21.
XX P-PSDB; AAU85524.
XX Novel polynucleotide encoding a lung tumour polypeptide useful for
XX stimulating and/or expanding T cells specific for a tumour protein -
PT
XX Example 1; SEQ ID No 790; 223pp; English.
XX
XX The invention describes an isolated polynucleotide and polypeptide
XX useful for stimulating and/or expanding T cells specific for a tumour
XX protein for determining the presence of a cancer in a patient. A
XX composition containing the polynucleotide and/or polypeptide is useful
XX for treating a lung cancer in a patient. The polypeptide is useful for
XX removing tumour cells from a biological sample. The polynucleotide is
XX also useful as probe or primer to detect the level of mRNA encoding a
XX tumour protein. This sequence encodes a lung tumour associated protein
XX or protein fragment, described in the method of the invention.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 457 BP; 146 A; 105 C; 121 G; 85 T; 0 other;
SQ

Query Match 30.5%; Score 204; DB 24; Length 457;
Best Local Similarity 77.4%; Pred. No. 1.4e-51;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
XX 319 GTCCAGGAAACCATGTAGAGAGCCCTTAAGAGAGAACCCACCTAAAGTCGGAA 378
DB 120 GTCCAGGAGCCCGAGTAATGGAGAGCCCAAAAGAGAGACGAGCTGAAAGTCGGGA 179
XX 379 TCTACACCTGTATCAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGTGCCTGA 438
DB 180 TCTACACCTGGCAGCAGACAGAGAGATCAGATACAGTCCAGTCCAGTGGCGGA 239
XX 439 CTGGAAGCCGATCTCCAGGAGCTATGTGAGCAAAAGACTGGGGATGGATGTGAAGGT-G 497
DB 240 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGG 299
XX 498 GTACTGATCTCAAGGGGAGATTTCTACCAAAAGCAGAGACTTTAAATGCCAGAGAGCAG 557
DB 300 GTTCCGGCTCAAGGTGAAGATAATACCTTAAGAGAGACATCTGTAATGCCAGAGCAG 359
XX 558 GTCAAGGGAAATCACAGGTTTAAAGGAAGATAAGCTGAAACACACAAACTGTTTTATA 617
DB 360 GTGAAGAGCAACCAAGTTTAAATGAAGACAAGCTGAAACCAACGCAAGCTGGTTTTATA 419
XX 618 TTAGATATTTTACTTTAAGAGCTCTTAATAAATTTT 653
DB 420 TTAGATATTTGACTTAAACTATCTCAATAAAGTTTT 455


```

PR 06-MAR-2000; 2000US-0519642.
PR 22-MAR-2000; 2000US-0533077.
PR 10-APR-2000; 2000US-0546259.
PR 27-APR-2000; 2000US-0560406.
PR 05-JUN-2000; 2000US-0589184.
PR 11-JUL-2000; 2000US-0614124.
PR 29-AUG-2000; 2000US-0651563.
PR 08-SEP-2000; 2000US-0658824.
PR 26-SEP-2000; 2000US-0671325.
PR 06-OCT-2000; 2000US-0677419.
PR 30-OCT-2000; 2000US-0702705.
PR 13-DEC-2000; 2000US-0736457.
PR 03-MAY-2001; 2001US-0849626.
XX
XX (CORI-) CORIYA CORP.
XX
XX Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW, Durham M;
XX Carter D, Fanger GR, Vedvick TS, Bangur CS, McNabb A;
XX
XX WPI; 2003-328427/31.
XX
XX New polynucleotide, useful for preparing a composition for treating or
XX inhibiting development of cancer, e.g. lung cancer -
XX
XX Example 1; SEQ ID NO 790; 82pp; English.
XX
XX The invention describes an isolated polynucleotide comprising one of 32
XX sequences, complement or degenerate variants of them. The polynucleotide
XX is useful for preparing a composition e.g. a vaccine or for gene therapy,
XX for treating or inhibiting development of cancer, e.g. lung cancer.
XX This sequence represents a polynucleotide associated with the
XX compositions and methods for the therapy and diagnosis of lung cancer.
XX
XX SQ Sequence 457 BP; 146 A; 105 C; 121 G; 85 T; 0 other;
XX
XX Query Match 30.5%; Score 204; DB 25; Length 457;
XX Best Local Similarity 77.4%; Pred. No. 1.4e-51;
XX Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
XX
XX QY 319 GTCCAGGAGCCACCTGATGAGAGGCTTAAAGAGAGAGAAACCCACCTAAAGTCCGAA 378
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 120 GTCCAGGAGCCACCTGATGAGAGGCTTAAAGAGAGAGAAACCCACCTGAAAGTCCGGA 179
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 379 TCCTACACTGATCAGAGAGAGAGATGATCAGGCTGACCTGAGATTCAGTGCCTGA 438
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 180 TCCTACACTGCGGACGACAGAGAGATGATCAGGCTGACCTGAGATTCAGTGCCTGA 239
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 439 COTGAAGCGATCTCCAGGAGCTATGTCAGACAAAGCTGGGATGATGTGAAGT-G 497
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 240 CATGAAGGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGATAAATCTGATTGG 299
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 498 GTACTGATGTCAGGGGAGATTTTACCAAGAGAGAGACATTTTAAATGCCAGAGCAG 557
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 300 GTTCCGGGCTCAAGTGAAGATAATACCTAAAGAGGACACTGTAAATGTCAGAGCAG 359
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 558 GTGAAGGAATACAGGTTTAAAGGAGATGATGAAACACACAAACCTGTTTTTATA 617
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 360 GTGAAGAGACACACAGTTTAAATGAGAGACAGCTGAAACACGCGAGCTGTTTTATA 419
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 618 TTAGATATTACTTTTAAAGAGCTTTAATAAATTTT 653
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 420 TTAGATATTGACTTAAACTACTTCTCAATAAAGTTTT 455
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX RESULT 23
XX ID ABA92217 standard; cDNA; 463 BP.
XX AC ABA92217;
XX
XX DT 06-JUN-2002 (first entry)
XX
XX Melanoma metastasis X-linked gene 1 (MMX-1) cDNA.

```

```

XX MMX-1; melanoma metastasis X-linked gene 1; human; tumour;
XX mammary carcinoma; breast cancer; marker; cancer/testis antigen;
XX CTA; chromosome Xp11.21-Xp11.22; diagnosis; gene therapy; gene; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 126..371
XX /tag= a
XX /product= "MMX-1_polypeptide"
XX /polya_signal 434..439
XX /tag= b
XX primer_bind complement (48..68)
XX /tag= c
XX primer_bind 385..404
XX /tag= d
XX
XX EP1179589-A1.
XX
XX 13-FEB-2002.
XX
XX 08-AUG-2000; 2000EP-0116253.
XX
XX 08-AUG-2000; 2000EP-0116253.
XX (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX
XX Not given;
XX
XX WPI; 2002-229925/29.
XX P-PSDB; AAW51134.
XX
XX New human cancer/testis antigen polypeptide, termed melanoma metastasis
XX X-linked gene 1 polypeptide, useful as specific marker of metastatic
XX tumor cells -
XX
XX Claim 1; Fig 1; 30pp; English.
XX
XX The present sequence is that of novel human melanoma metastasis
XX X-linked gene 1 (MMX-1) cDNA. The cDNA was identified following
XX suppression subtractive hybridisation analysis comparing common
XX naevi and melanoma metastases. MMX-1 expression is upregulated in
XX metastatic tumour cells and is induced during tumour expression
XX and/or metastasis, especially in malignant melanoma and mammary
XX carcinoma cells. MMX-1 is a specific marker of metastatic cancer
XX T-cells but is not presented in an MHC Class I complex to cytotoxic
XX cells expressing MMX-1 are testis cells which do not present
XX antigens in an MHC Class I context. The function of the MMX-1
XX gene is to promote loss of contact inhibition and anchorage
XX dependence in tumour cells and to promote other essential steps of
XX the metastatic cascade. Expression of the MMX-1 gene correlates
XX with a more aggressive behaviour of the tumour cells and also with
XX the potential for formation of metastasis. The expression profile
XX of MMX-1 in normal tissues, tumour cell lines and tumour samples
XX places the gene in the group of cancer/testis antigens (CTA). The
XX MMX-1 gene has been identified, spanning about 5 kb and consisting
XX of 4 exons and 3 introns, 2 of the introns being in the coding
XX region. The gene is localised on the X chromosome (Xp11.21-Xp11.22)
XX between markers DXS1204 and DXS1199, next to the CTA families of
XX SSX, GAGE and MAGE-D. The invention provides methods for detecting
XX the metastatic potential of melanoma and mammary carcinoma cells,
XX and a process for determining whether a test sample of human cells
XX has tumour progression potential, indicated by a greater amount of
XX MMX-1 nucleic acid than cancer cell samples free of metastasis.
XX Also provided are methods for identifying antagonists of MMX-1 or
XX inhibitors of MMX-1 expression (e.g. antisense nucleotides) that
XX can be used to inhibit tumour progression or metastasis and cause
XX apoptosis of tumour cells in vivo. The MMX-1 gene and protein (see
XX AAW51134) can also be used to identify and design drugs which
XX interfere with the proliferation and dissemination of tumour cells.
XX

```


PR 15-OCT-1999; 99US-0419356.
PR 17-DEC-1999; 99US-0466867.
PR 30-DEC-1999; 99US-0476300.
PR 06-MAR-2000; 2000US-0519642.
PR 22-MAR-2000; 2000US-0533077.
PR 10-APR-2000; 2000US-0546259.
PR 27-APR-2000; 2000US-0560406.
PR 05-JUN-2000; 2000US-0589184.
XX (CORI-) CORIXA CORP.
XX Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
PI Retter MW, Mannion J;
XX WPI: 2001-071489/08.
XX Lung tumor-associated proteins and the nucleic acids that encode them,
PT useful for preventing, diagnosing and treating lung cancer -
XX
PS Claim 4; Page 426; 436pp; English.
XX
CC The present invention describes immunogenic portions of lung tumour-
CC associated proteins (I) and the nucleic acids (NAs) that encode them.
CC (I) have cytostatic activity and can be used in gene therapy, antisense
CC inhibition and in vaccines. The NAs and the lung tumour-associated
CC proteins they encode may be used in the prevention, treatment and
CC diagnosis of diseases associated with their inappropriate expression,
CC especially lung cancers. For example, the NAs may be administered to
CC treat diseases by rectifying mutations or deletions in a patient's genome
CC that affect the activity of the protein by expressing inactive proteins
CC or to supplement the patients own production of (I). Additionally, the
CC NAs may be used to produce the lung-tumour associated protein, according
CC to standard recombinant DNA methodology. Conversely, antisense NA
CC molecules may be administered to down regulate protein expression by
CC binding with the cells own genes and preventing their expression. The NA
CC and complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar NA sequences in
CC samples, and hence which patients may be in need of treatment for lung
CC cancer. The (I) may be used as antigens in the production of antibodies
CC and in assays to identify modulators (agonists and antagonists) of the
CC expression and activity of the protein. AAF68083 to AAF68878 and
CC AAB76848 to AAB76878 represent human lung tumour protein related
CC nucleotide and protein sequences which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 781 BP; 230 A; 181 C; 234 G; 136 T; 0 other;

Query Match 30.5%; Score 204; DB 22; Length 781;
Best Local Similarity 77.4%; Pred. No. 1.8e-51;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGAGCCCACTGATGAAGAGCCTTAAAGAGAGAAACCCACCTAAAGTGGAA 378
DB 422 GTCCAGGAGCCCACTGATGAAGAGCCTTAAAGAGAGAAACCCACCTAAAGTGGAA 481
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTGTCAGCTGAGATTCAAGTCTGA 438
DB 482 TCCTACACCTGGGAGAGAGAGAGAGATGATCAGGTGTCAGCTGAGATTCAAGTCTGA 541
QY 439 CCTGGAAGCCGATCTCCAGGAGCTATGTCACACAAAGACTGGGGATGGATGAGGT-G 497
DB 542 CATGAGAGGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGATTAATCTGATTTGG 601
QY 498 GTACTGATGTCAGGGAGAGATTCTACAAAGAGAGAGACTTTAAATGCCAGAGCAG 557
DB 602 GTTCGGGCTCAGGTGAAGATAATACCTAAAGAGGAGAACACTGTAAATGCCAGAGCAG 661
QY 558 GTGAAGGGAATACAGGTTTAAAGGAGATAAGCTGAACACACAACTGTTTTTATA 617
DB 662 GTGAAGAGACACCAAGTTTAAATGAGAGACAGCTGAACACAGCAAGCTGTTTTATA 721
QY 618 TTAGATATTTTAAAGAGCTTTTAAATATTTT 653

DB 722 TTAGATATTTGACTTAAACTATCTCAATAAAGTTTT 757
RESULT 29
ABK38772
ID ABK38772 standard; cDNA; 781 BP.
XX
AC ABK38772;
XX
XX 21-MAY-2002 (first entry)
XX
DE cDNA encoding L552S lung tumour protein.
XX
KW Lung tumour; cancer; T cell; immune response stimulator;
XX cytotstatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200204514-A2.
XX
XX 17-JAN-2002.
XX
XX 10-JUL-2001; 2001WO-US22058.
XX
XX 11-JUL-2000; 2000US-0614124.
XX 23-AUG-2000; 2000US-0651563.
XX 08-SEP-2000; 2000US-0658824.
XX 26-SEP-2000; 2000US-0671325.
XX 06-OCT-2000; 2000US-0677419.
XX 30-OCT-2000; 2000US-0702705.
XX 13-DEC-2000; 2000US-0736457.
XX 03-MAY-2001; 2001US-0849626.
XX (CORI-) CORIXA CORP.
XX Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
PI Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS;
PI McNabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;
XX
XX WPI: 2002-164634/21.
XX P-PSDB; AAU55651.
XX
PT Novel polynucleotide encoding a lung tumour polypeptide useful for
PT stimulating and/or expanding T cells specific for a tumour protein -
XX
XX Example 1; SEQ ID No 808; 223pp; English.
XX
CC The invention describes an isolated polynucleotide and polypeptide
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein for determining the presence of a cancer in a patient. A
CC composition containing the polynucleotide and/or polypeptide is useful
CC for treating a lung cancer in a patient. The polypeptide is useful for
CC removing tumour cells from a biological sample. The polynucleotide is
CC also useful as probe or primer to detect the level of mRNA encoding a
CC tumour protein. This sequence encodes a lung tumour associated protein
CC or protein fragment, described in the method of the invention.
CC Note: the sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 781 BP; 230 A; 181 C; 234 G; 136 T; 0 other;

Query Match 30.5%; Score 204; DB 24; Length 781;
Best Local Similarity 77.4%; Pred. No. 1.8e-51;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGAGCCCACTGATGAAGAGCCTTAAAGAGAGAAACCCACCTAAAGTGGAA 378
DB 422 GTCCAGGAGCCCACTGATGAAGAGCCTTAAAGAGAGAAACCCACCTAAAGTGGAA 481
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTGTCAGCTGAGATTCAAGTCTGA 438

```
Db 482 TCCTACACCTGGCGCAGCAGACAGAGAAGATCAGGATACAGCTGAGATCCAGTGGCGGA 541
Qy 439 CCTGGAGCCGATCTCCAGGAGCTATGTGACAGAAAGACTGGGATGGATGTGAAGT-G 497
Db 542 CATGGAAGGTGATCTGAGAGCTGCATCAGTCAACACCGGGATAAATCTGATTTGG 601
Qy 498 GTACTGATGTCAAGGGGAAGATTCTACCAAGCAGAGCAGCTTTAAATGCCAGAGCAG 557
Db 602 GTTCGGCGTCAAGGTGAAGATAATACCTAAAGAGGACACTGTANAATGCCAGAGCAG 661
Qy 558 GTGAAGGAAATCACAGGTTTAAAGGAGATAAGCTGAACAAACACAACTGTTTTTATA 617
Db 662 GTGAAGACACACCAAGTTTAAATGAAGACAGCTGAACAAACGCAAGCTGTTTTATA 721
Qy 618 TTAGATATTTTACTTTAAAGAGCTTTAATAAATTTT 653
Db 722 TTAGATATTTGACTTAAACTATCTCATATAAAGTTTT 757

RESULT 30
ACAA1101
ID ACA11101 standard; cDNA; 781 BP.
AC ACAA11101;
XX
XX
XX 05-JUN-2003 (first entry)
DT Human lung cancer-associated full length cDNA L5528.
DE Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
KW T cell expansion; CD4; CD8.
XX
XX Homo sapiens.
OS
XX
XX US2002197669-A1.
XX
XX 26-DEC-2002.
XX
XX 03-MAY-2001; 2001US-0849626.
XX
XX 13-DEC-2000; 2000US-0736457.
XX
XX (BANG/) BANGUR C S.
PA (FANG/) FANGER G R.
PA (WANG/) WANG A.
PA (WANG/) WANG T.
PA (SWIT/) SWITZER A P.
PA (MCNE/) MCNEILL P D.
PA (CLAP/) CLAPPER J D.
XX
XX Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;
PI Clapper JD;
XX
XX WPI; 2003-352750/33.
DR P-PSDB; ABU69504.
XX
XX Novel lung cancer polynucleotide encoding lung cancer protein, useful
PT for detecting the presence of lung cancer in a patient, and in
PT pharmaceutical compositions, e.g. vaccines, for treating lung cancer -
XX
XX Example 1; Page -; 72pp; English.
XX
XX The invention relates to a polynucleotide encoding a lung tumour protein,
XX comprising a sequence selected from any of the 14 sequences
XX mentioned in the specification, or a sequence (S2) mentioned in
XX the specification, complement of S1, sequences consisting of at least 20
XX contiguous residues of S1, sequences that hybridise to S1, sequences
XX having 75%, preferably 90%, identity to S1, or degenerate variants of
XX S1. Also included are an isolated polypeptide (comprising a sequence (S3)
XX selected from any one of the 4 amino acid sequences mentioned in the
XX specification, a sequence encoded by the polynucleotide, or sequences
XX having at least 70%, preferably 90%, identity to a sequence encoded by
XX the polynucleotide), an expression vector comprising the polynucleotide
```

```
CC operably linked to an expression control sequence, a host cell
CC transformed or transfected with the vector, an isolated antibody (or its
CC antigen-binding fragment) that specifically binds to the polypeptide,
CC detecting the presence of a cancer in a patient, a fusion protein
CC comprising the polypeptide, an oligonucleotide that hybridises to
CC si under moderately stringent conditions, stimulating and/or expanding T
CC cells specific for a tumour protein (comprising contacting T cells with
CC the polynucleotide, protein or antigen-presenting cells, under conditions
CC and for a time sufficient to permit the stimulation and/or expansion of T
CC cells) and inhibiting the development of a cancer in a patient (by
CC incubating CD4+ and/or CD8+ T cells isolated from a patient with the
CC polynucleotide, protein or antigen presenting cells that express the
CC polynucleotide, such that T cells proliferate, administering to the
CC patient an effective amount of the proliferated T cells, and thus
CC inhibiting the development of a cancer in the patient. The
CC polynucleotide, protein and cells are useful in a composition for
CC stimulating an immune response in a patient, and for treating a cancer in
CC a patient (particularly lung cancer). The oligonucleotide is useful for
CC determining the presence of a cancer in a patient. The protein and
CC oligonucleotides are useful in pharmaceutical compositions, e.g.
CC vaccines. The polynucleotide is also useful as a probe or primer for
CC nucleic acid hybridisation, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of tumour polypeptides and
CC proteins in tumour cells. An amplified portion of the polynucleotide is
CC useful for isolating a full-length gene from a suitable library.
CC The present sequence is a cDNA (full length, extended or partial)
CC isolated from a library derived from lung tumour/cancer cells.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from the USPTO
CC at seqdata.uspto.gov/sequence.html?DocId=20020197669.
XX
XX SQ Sequence 781 BP; 230 A; 181 C; 234 G; 136 T; 0 other;
Query Match 30.5%; Score 204; DB 25; Length 781;
Best Local Similarity 77.4%; Pred. No. 1.8e-51;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
Qy 319 GTCCAGGAAACCCACCTGATGAGAGAGCTAAAGAGAGAAACCCACCTAAGAGTCGGAA 378
Db 422 GTCCAGGAGCCCAAGTAATGGAGAGCCCAAGAGAGAAACCCAGCAGCTGAAGTCGGGA 481
Qy 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGCTGCAGCTGAGATTCAAGTGCCTGA 438
Db 482 TCCTACACCTGGCAGCAGACAGAGAGATCAGGATACAGCTGAGATCCAGTGGCGGA 541
Qy 439 CTGGAAGCCGATCTCCAGAGCTATGTGACAGAAAGACTGGGGATGGATGTGAAGT-G 497
Db 542 CATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGGATAAATCTGGATTGG 601
Qy 498 GTACTGATGTCAAGGGGAAGATTCTACCAAGCAGAGCAGACTTTAAATGCCAGAGCAG 557
Db 602 GTTCCGCGCTCAAGGTGAAGATAATACCTAAGAGGACACTGTANAATGCCAGAGCAG 661
Qy 558 GTGAAGGGAATCACAGGTTTAAAGGAGATAGCTGAACAAACACAACTGTTTTTATA 617
Db 662 GTGAAGAGCAGCAACCAAGTTTAAATGAAGACAGCTGAACAAACGCAAGCTGTTTTATA 721
Qy 618 TTAGATATTTTACTTTAAAGAGCTTTAATAAATTTT 653
Db 722 TTAGATATTTGACTTAAACTATCTCATATAAAGTTTT 757

RESULT 31
ACAA02287
ID ACA02287 standard; cDNA; 781 BP.
XX
XX ACA02287;
AC ACA02287;
XX
XX 22-MAY-2003 (first entry)
DT Lung cancer therapyand diagnosis associated cDNA #776.
XX
XX
```



```
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 750 BP; 225 A; 164 C; 203 G; 158 T; 0 other;

Query Match 27.5%; Score 193.8; DB 23; Length 750;
Best Local Similarity 86.4%; Pred. No. 2.5e-45;
Matches 203; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 323 CAGGACCCACTGATGAGAGCCCTAAAGAGAGAAACACCCACTAAAGTCGGAATCCT 382
DB 208 CAGGAGCCCATGTGCGCAGAGCCCTCAACAGAGAAACACCACTGAAAGTCAGGATCAT 267
QY 383 ACACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTSCCTGACCTG 442
DB 268 ACACCTGCTCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTSCCTTAACCTG 327
QY 443 GAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGATGTGAGGTGCTACT 502
DB 328 GAAGCTGATCTCCAGGAGCTGTCTCAGTCAAGACTGGGGATGAATGCGGAGATAGTCCT 387
QY 503 GATGTCAAGGGGAGATTCTACCAAAAGCAGAGCACATTTAAATGCCAGAGCAG 557
DB 388 GATGTCCAGGGGAGATTCTGCCAAATCAGAGCAATTTAAATGCCAGAGGAG 442

RESULT 35
ABZ20463
ID ABZ20463 standard; cDNA; 530 BP.
AC ABZ20463;
XX
DT 23-JAN-2003 (first entry)
XX
DE GAGE-2 full length cDNA sequence SEQ ID NO:2890.
XX
KW Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
KW immune response; virology; immunology; microbiology; molecular biology;
KW recombinant DNA technology; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200278516-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US10421.
XX
PR 30-MAR-2001; 2001US-280255P.
PR 28-AUG-2001; 2001US-315563P.
PR 09-JAN-2002; 2002US-347313P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang S, Bangur CS, Gaiger A;
XX
DR WPI; 2003-058387/05.
DR P-PSDB; ABP54447.
XX
PT New immunogenic polynucleotides or polypeptides useful for diagnosing,
PT preventing and treating cancer expressing CT or CP mRNA antigens, and
PT in virology, immunology, microbiology, molecular biology and
PT recombinant DNA techniques
XX
PS Claim 1; SEQ ID 2890; 207pp; English.
XX
CC ABQ17575 to ABQ20506 represent isolated polynucleotide (i) sequences, and
CC ABP54446 to ABP54472 represent protein (ii) sequences, from the present
CC invention. (i) and (ii) have cytostatic activity and can be used in gene
CC therapy and vaccines. (i), (ii), antibodies and compositions from the
CC present invention are useful for diagnosing, preventing and treating
```

```
CC cancer, which expresses CT or CP mRNA antigens. They are useful for
CC stimulating immune response. They can also be useful in virology,
CC immunology, microbiology, molecular biology and recombinant DNA
CC techniques.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 530 BP; 152 A; 116 C; 155 G; 107 T; 0 other;

Query Match 27.1%; Score 181.2; DB 25; Length 530;
Best Local Similarity 67.4%; Pred. No. 1.3e-44;
Matches 287; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

QY 217 GTGAATATGATGTGGCGAGGAAGATCAACATATAGGCTAGGCCCAAGAGAGTTTACA 276
DB 77 GTGAATATGATGTGGCGAGGAAGATCGACCTATCGGCTAGACCAAGACGCTACGTAGA 136
QY 277 GCCTCTCTGAGCTGATTGGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCACTGA 336
DB 137 GCCTCTCTGAAATGATTGGCCCTATGGGCC---CGAGCAGTTTCAGTGATGAAGTGAACC 193
QY 337 TGAAGAGCCTAAAGAGAGAAAACACCCACTAAAAGTGGGAATCCTACACCTGATCAGNA 396
DB 194 AGCAACACCTGAAGAGAGGGGAAACAGCAACTCAAGTCAGGATCCTGCAGCTGCTCAGGA 253
QY 397 GAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAGCCGATCTCCA 456
DB 254 GGGAGAGGATG---AGGGAGCATCTGCAGTCAAGGCCCAAGCTGAAGCTCATAGCCA 310
QY 457 GGAGCTATCTCAGACAAAGACTGGGGATGGATGTGAAGTGTGTACTGTATGTCAAGGGGAA 516
DB 311 GGAACAGGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGCTCTGTATGGCAGGAGAT 370
QY 517 GATTCTACCAAAAGCAGAGCACCTTTAAATGCCAGAGCAGCTGAAGGGAATCACAGT 576
DB 371 GGACCCGCCAATCCAGAGAGGTGAACGCTGAAGAGGTGAAGCAATCACAGTG 430
QY 577 TTAAGAGAGATAAGCTGAACAAACAAACAACTGTGTTTTATATAGATATTTTACTTTAA 636
DB 431 TTAAGAGAGACACGTTGAATGATGACAGGCTGCTCTATGTTGGAATTTTGTTCATTAA 490
QY 637 GAGTCT 642
DB 491 AATTCT 496

RESULT 36
AAV18717
ID AAV18717 standard; cDNA; 535 BP.
XX
AC AAV18717;
XX
DT 30-JUL-1998 (first entry)
XX
DE cDNA encoding GAGE-2 tumour rejection antigen precursor.
XX
KW GAGE tumour rejection antigen precursor; TRAP; tumour;
KW diagnosis; melanoma; antigen; cytolytic T cell clone proliferation;
KW HLA-typing assay; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
XX CDS 81..431
XX /*tag= a
XX /*transl_except= (pos:192..194, aa:Ala)
XX /*transl_except= (pos:195..197, aa:Thr)
XX
XX MO9749417-A1.
XX
XX 31-DEC-1997.
XX
```



```

PF 23-JUN-1997; 97WO-US10850.
XX
XX PR 24-JUN-1996; 96US-0669161.
XX
XX PA (LUDW-) LUDWIG INST CANCER RES.
XX
XX PI Boon-Falleur T, Debacker O, Van Den Eynde B;
XX
XX DR WPI; 1998-076905/07.
XX P-PSDB; AAW47599.
XX
XX PT Isolated nucleic acid encoding GAGE tumour rejection antigen
XX precursor - processed by HLA-Cw6 molecules into peptides, useful to
XX diagnose melanomas
XX
XX PS Example 13; Fig 4; 60pp; English.
XX
XX CC The present sequence encodes a GAGE-2 tumour rejection antigen
XX precursor (TRAP). The protein is expressed in a number of tumours. In
XX contrast the only normal tissue which expresses GAGE TRAP protein is
XX testis. Several GAGE TRAPs have been identified (see AAV18717-21). The
XX major difference between these proteins and GAGE-1 (AAV05540) is the
XX absence of a stretch of 143 bases located at position 379 to 521 of the
XX GAGE-1 TRAP sequence. The rest of the sequences show mismatches at
XX various position, with the exception of GAGE-3 whose 5' end is totally
XX different from the other GAGE cDNAs for the first 112 bases. This
XX region of GAGE-3 cDNA contains a long repeat and a hairpin structure.
XX The antigens can be used to diagnose melanomas, characterised by
XX expression of a TRAP or presentation of a tumour rejection antigen.
XX Antigens shed into blood or urine can be observed and then used to
XX confirm a diagnosis of melanoma using cytolytic T cell clone
XX proliferation methodologies. Other uses for the processed peptides,
XX include HLA-typing assays for, e.g. skin graft or organ transplants.
XX
XX SQ Sequence 535 BP; 158 A; 112 C; 156 G; 109 T; 0 other;

Query Match 27.1%; Score 181.2; DB 19; Length 535;
Best Local Similarity 67.4%; Pred. No. 1.3e-44;
Matches 287; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

QY 217 GTGAAATATGAGTTGGCGAGGAAGATCAACATATAGGCTAGGCCAAGAAAGTTTACA 276
DB 74 GTGAAATATGAGTTGGCGAGGAAGATCGACCTATCGGCTAGCAAGACGCTACGTAGA 133

QY 277 GCCTCTGAGCTGATTGGGGCTATGCTTACTGGCTCCCTTTGTCCCGAGACCCACTGA 336
DB 134 GCCTCTGAAATGATTGGGCCCTATGGGCC---CGAGCAGTTTCAGTGATGAGTGGAAACC 190

QY 337 TGAAGAGCCTTAAGAAAGAGAAACCCACCTAAAGTGGAAATCCTACACCTGATCAGAA 396
DB 191 AGAGCCTCTGAAGAAAGGGGACCCAGCAACTCAAGTCAGGATCCTGCAGCTGCTCAGGA 250

QY 397 GAGAGAGATGATCAGGTTGCGAGTCAAGATTCAGTGCTGACCTGGAGCGGATCTCCA 456
DB 251 GGGAGAGGATG---AGGAGCAGTCTGAGGTCAAGGGCCGAAGCCTGAAGCTGAGAGCCA 307

QY 457 GGAGCTATGTTCAGACAAAGACTGGGGATGGATGTGAAGGTGGTACTGATGTCAAGGGAA 516
DB 308 GGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGTCCTGATGGCGCAGGAT 367

QY 517 GATTCTACCAAGAGCAGGACTTTAAATGCCAGNAGCAGGTGAAGGAATTCAGGT 576
DB 368 GGACCCGCCCAATCCAGAGAGGATGAAGACGCGCTGAAGAGGTGAAGCAATCAGTG 427

QY 577 TTAAGGAAGATAGCTGAAGAACAAACAACTGTTTTATATTAGATATTTTACTTTAAA 636
DB 428 TTAAGGAAGACATGTTGAATGATGATGAGGCTGCTCTCTATGTTGGAATTTGTTATTAA 487

QY 637 GAGTCT 642
DB 488 AATTCT 493

```

```

RESULT 37
AAK90519
ID AAK90519 standard; cDNA; 538 BP.
XX
XX AC AAK90519;
XX
XX DT 30-SEP-1999 (first entry)
XX
XX DE GAGE-2 tumour rejection antigen clone nucleotide sequence.
XX
XX KW Human leukocyte antigen; HLA-A29; tumour rejection antigen;
XX detection; therapy; pathological condition; cancer; CTL;
XX cytolytic T lymphocyte; GAGE; ss.
XX
XX OS Homo sapiens.
XX
XX PN W09937665-A1.
XX
XX PD 29-JUL-1999.
XX
XX PF 12-JAN-1999; 99WO-US00775.
XX
XX PR 23-JAN-1998; 98US-0012818.
XX
XX PA (LUDW-) LUDWIG INST CANCER RES.
XX
XX PI Boon-Falleur T, Debacker O, Van Den Eynde B, Van Der Bruggen P;
XX WPI; 1999-469111/39.
XX
XX PT New isolated peptides which bind to HLA-A29 molecules, which are
XX tumour rejection antigens used for detection and therapy of
XX pathological conditions, e.g. cancer
XX
XX PS Example 13; Fig 4; 62pp; English.
XX
XX CC The present invention describes peptides which bind to human leukocyte
XX antigen (HLA) A29 (HLA-A29) molecules. The peptides are processed into
XX tumour rejection antigens. They can be used for detecting cytolytic T
XX lymphocytes (CTLs) in pathological conditions such as cancer and in
XX HLA-typing assays. Complexes of HLA-A29 molecules and the peptides can
XX be used for stimulating CTLs in vivo. The present sequence represents
XX a GAGE tumour rejection antigen clone, from an example from the present
XX invention.
XX
XX SQ Sequence 538 BP; 160 A; 116 C; 155 G; 107 T; 0 other;

Query Match 27.1%; Score 181.2; DB 20; Length 538;
Best Local Similarity 67.4%; Pred. No. 1.3e-44;
Matches 287; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

QY 217 GTGAAATATGAGTTGGCGAGGAAGATCAACATATAGGCTAGGCCAAGAAAGTTTACA 276
DB 77 GTGAAATATGAGTTGGCGAGGAAGATCGACCTATCGGCTAGCAAGACGCTACGTAGA 136

QY 277 GCCTCTGAGCTGATTGGGGCTATGCTTACTGGCTCCCTTTGTCCAGAAACCCACTGA 336
DB 137 GCCTCTGAAATGATTGGGGCTATGGGCC---CGAGCAGTTTCAGTGATGGAAGTGAAC 193

QY 337 TGAAGAGCCTTAAGAAAGAGAAACCCACCTAAAGTCCGAAATCCTACACCTGATCAGAA 396
DB 194 AGCAACACCTTGAAAGAGGGGAAACAGCACTCAAGCTCAGGATCTTCGAGCTGCTCAGGA 253

QY 397 GAGAGAAAGATGATCAGGTTGCGAGTTCAGATTCAAGTTCCTGAGCTCGAAGCCGATCTCA 456
DB 254 GGGAGAGGATG---AGGAGCAGTCTGCGGTCAAGGGCCGAAAGCCTGAAGCTCATAGCCA 310

QY 457 GGAGCTATGTTCAGACAAAGACTGGGGATGGATGAGTGGTGGTACTGATGTCAAGGGAA 516
DB 311 GGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGGTTCCTGATGGCGAGGAT 370

QY 517 GATTCTACCAAGAGCAGGACTTTAAATGCCAGAAAGCAGGTGAAGGGAAATCAGAGT 576

```

Db 371 GGACCGCCAAATCCAGAGGAGGTGAAGAACCCCTGAAGAGGTGAAGAGCAATCACAGTG 430

Qy 577 TTAAGGAGATAGCTGAACACACAAACTGTTTATATATAGATATTTTACTTTAA 636

Db 431 TTAAGGAGACACGTTGAATGATGCGAGGCTGCTCTATGTTGGAATTTGTTCAATA 490

Qy 637 GAGTCT 642

Db 491 AATCT 496

RESULT 38

AAF68152

ID AAF68152 standard; cDNA; 479 BP.

XX

AC AAF68152;

XX

DT 12-APR-2001 (first entry)

XX

DE Human lung tumour protein related nucleotide sequence SEQ ID NO:70.

XX

KW Human; lung cancer; lung tumour; lung tumour protein; gene therapy;

KW lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;

KW cytostatic; antisense inhibition; ss.

XX

OS Homo sapiens.

XX

FN WO200100828-A2.

XX

PD 04-JAN-2001.

XX

PF 30-JUN-2000; 2000WO-US18061.

XX

PR 30-JUN-1999; 99US-0346492.

PR 15-OCT-1999; 99US-0419356.

PR 17-DEC-1999; 99US-0456867.

PR 30-DEC-1999; 99US-0476300.

PR 06-MAR-2000; 2000US-0519642.

PR 22-MAR-2000; 2000US-0532077.

PR 10-APR-2000; 2000US-0546259.

PR 27-APR-2000; 2000US-0560406.

PR 05-JUN-2000; 2000US-0589184.

XX

PA (CORI-) CORIXA CORP.

XX

PI Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;

PI Retter MW, Mannion J;

XX

WPI; 2001-071488/08.

XX

Lung tumor-associated proteins and the nucleic acids that encode them,

PT useful for preventing, diagnosing and treating lung cancer -

XX

PS Claim 4; Page 174; 436pp; English.

XX

CC The present invention describes immunogenic portions of lung tumour-

CC associated proteins (I) and the nucleic acids (NAs) that encode them.

CC (I) have cytostatic activity and can be used in gene therapy, antisense

CC inhibition and in vaccines. The NAs and the lung tumour-associated

CC proteins they encode may be used in the prevention, treatment and

CC diagnosis of diseases associated with their inappropriate expression,

CC especially lung cancers. For example, the NAs may be administered to

CC treat diseases by rectifying mutations or deletions in a patient's genome

CC that affect the activity of the protein by expressing inactive proteins

CC or to supplement the patients own production of (I). Additionally, the

CC NAs may be used to produce the lung-tumour associated protein, according

CC to standard recombinant DNA methodology. Conversely, antisense NA

CC molecules may be administered to down regulate protein expression by

CC binding with the cells own genes and preventing their expression. The NA

CC and complementary sequences may also be used as DNA probes in diagnostic

CC assays to detect and quantitate the presence of similar NA sequences in

CC samples, and hence which patients may be in need of treatment for lung

CC cancer. The (I) may be used as antigens in the production of antibodies

CC

CC and in assays to identify modulators (agonists and antagonists) of the

CC expression and activity of the protein. AAF68083 to AAF68878 and

CC AAB76848 to AAB76878 represent human lung tumour protein related

CC nucleotide and protein sequences which are used in the exemplification

CC of the present invention.

XX

SQ Sequence 479 BP; 163 A; 107 C; 125 G; 84 T; 0 other;

Query Match 26.9%; Score 180; DB 22; Length 479;

Best Local Similarity 74.0%; Pred. No. 2.8e-44;

Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;

Qy 319 GTCCAGGAACCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCTAAAGTCGGAA 378

Db 103 GTCCAGGAGCCAGTATGGAGAGCCCAAAAGAGAACCCAGCAGTGAAGTCGGGA 162

Qy 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTC----- 428

Db 163 TCCTACACCTGGGCAGCAGACAGAGAGATCAGGATACAGCTGAGATCCCGAGTCTGG 222

Qy 429 -----AAGTCTGACCTGGAAGCGGATCTCCAGAGCTATGTCAGACAAAGACTGGGG 482

Db 223 GAAGGGAATGCGGACATGGAAGTGATCTCCAAAGAGCTGCATCAGTCAACACCGGGG 282

Qy 483 ATGGATGTGAAGGT-GGTACTGATGTCAAGGGGAAGATTCTACCAAAAGCAGAGCACTTT 541

Db 283 ATAAATCTGGATTTGGGTTCCGCGCTCAAGGTGAAGATAATACCTAAAGAGGAACACTGT 342

Qy 542 AAAATGCCAGAGCAGGTGAAGGGAATACACAGGTTTAAAGAGATAGCTGAACACAC 501

Db 343 AAAATGCCAGAGCAGGTGAAGGGAATACACAGGTTTAAAGAGATAGCTGAACACAC 402

Qy 602 ACAAACTGTTTTATATAGATATTTTACTTTAAAGAGTCTTTAATAATTTTG 655

Db 403 GCAAGCTGTTTTATATAGGATATTTGACTTTAAACTATCTCAATAAAGTTTG 456

RESULT 39

ABK38063

ID ABK38063 standard; cDNA; 479 BP.

XX

AC ABK38063;

XX

DT 21-MAY-2002 (first entry)

XX

DE cDNA encoding clone #19106 (L547S) of lung tumour protein.

XX

KW Lung tumour; cancer; T cell; immune response stimulator;

KW cytostatic; gene; ss.

XX

OS Homo sapiens.

XX

FN WO200204514-A2.

XX

PD 17-JAN-2002.

XX

PF 10-JUL-2001; 2001WO-US22058.

XX

PR 11-JUL-2000; 2000US-0614124.

PR 23-AUG-2000; 2000US-0651563.

PR 08-SEP-2000; 2000US-0658824.

PR 26-SEP-2000; 2000US-0671325.

PR 06-OCT-2000; 2000US-0677419.

PR 30-OCT-2000; 2000US-0702705.

PR 13-DEC-2000; 2000US-0736457.

PR 03-MAY-2001; 2001US-0849626.

XX

(CORI-) CORIXA CORP.

PA

XX

PI Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;

PI Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS;

PI McNabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;

XX

DR WPI; 2002-164634/21.
XX Novel polynucleotide encoding a lung tumour polypeptide useful for
PT stimulating and/or expanding T cells specific for a tumour protein
XX
PS Example 1; SEQ ID No 70; 223pp; English.
XX
CC The invention describes an isolated polynucleotide and polypeptide
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein for determining the presence of a cancer in a patient. A
CC composition containing the polynucleotide and/or polypeptide is useful
CC for treating a lung cancer in a patient. The polypeptide is useful for
CC removing tumour cells from a biological sample. The polynucleotide is
CC also useful as probe or primer to detect the level of mRNA encoding a
CC tumour protein. This sequence encodes a lung tumour associated protein
CC or protein fragment, described in the method of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 479 BP; 163 A; 107 C; 125 G; 84 T; 0 other;
SQ
Query Match 26.9%; Score 180; DB 24; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
RESULT 40
ID ACA10392
XX ACA10392 standard; cDNA; 479 BP.
AC ACA10392;
XX
XX 05-JUN-2003 (first entry)
XX Human lung cancer-associated cDNA, L547D.
XX Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
KW T cell expansion; CD4; CD8.
XX
XX Homo sapiens.
XX
XX US2002197669-A1.
XX
XX 26-DEC-2002.
XX
XX 03-MAY-2001; 2001US-0849626.
XX

PR 13-DEC-2000; 2000US-0736457.
XX (BANG/) BANGUR C S.
PA (FANG/) FANGER G R.
PA (WANG/) WANG A.
PA (WANG/) WANG T.
PA (SWIT/) SWITZER A P.
PA (MCNE/) MCNEILL P D.
PA (CLAP/) CLAPPER J D.
XX
XX Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;
PI Clapper JD;
PI
XX WPI; 2003-352750/33.
XX
XX Novel lung cancer polynucleotide encoding lung cancer protein, useful
PT for detecting the presence of lung cancer in a patient, and in
PT pharmaceutical compositions, e.g. vaccines, for treating lung cancer
XX
XX Example 1; Page -; 72pp; English.
XX
XX The invention relates to a polynucleotide encoding a lung tumour protein,
CC comprising a sequence selected from any of the 14 sequences
CC mentioned in the specification, or a sequence (S2) mentioned in
CC specification, complement of S1, sequences consisting of at least 20
CC contiguous residues of S1, sequences that hybridise to S1, sequences
CC having 75%, preferably 90%, identity to S1, or degenerate variants of
CC S1. Also included are an isolated polypeptide (comprising a sequence (S3)
CC selected from any one of the 4 amino acid sequences mentioned in the
CC specification, a sequence encoded by the polynucleotide, or sequences
CC having at least 70%, preferably 90%, identity to a sequence encoded by
CC the polynucleotide), an expression vector comprising the polynucleotide
CC operably linked to an expression control sequence, a host cell
CC transformed or transfected with the vector, an isolated antibody (or its
CC antigen-binding fragment) that specifically binds to the polypeptide,
CC detecting the presence of a cancer in a patient, a fusion protein
CC comprising the polypeptide, an oligonucleotide that hybridises to
CC S1 under moderately stringent conditions, stimulating and/or expanding T
CC cells specific for a tumour protein (comprising contacting T cells with
CC the polynucleotide, protein or antigen-presenting cells, under conditions
CC and for a time sufficient to permit the stimulation and/or expansion of T
CC cells) and inhibiting the development of a cancer in a patient (by
CC incubating CD4+ and/or CD8+ T cells isolated from a patient with the
CC polynucleotide, protein or antigen presenting cells that express the
CC polynucleotide, such that T cells proliferate, administering to the
CC patient an effective amount of the proliferated T cells, and thus
CC inhibiting the development of a cancer in the patient. The
CC polynucleotide, protein and cells are useful in a composition for
CC stimulating an immune response in a patient, and for treating a cancer in
CC a patient (particularly lung cancer). The oligonucleotide is useful for
CC determining the presence of a cancer in a patient. The protein and
CC oligonucleotides are useful in pharmaceutical compositions, e.g.
CC vaccines. The polynucleotide is also useful as a probe or primer for
CC nucleic acid hybridisation, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of tumour polypeptides and
CC proteins in tumour cells. An amplified portion of the polynucleotide is
CC useful for isolating a full-length gene from a suitable library.
CC The present sequence is a cDNA (full length, extended or partial)
CC isolated from a library derived from lung tumour/cancer cells.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from the USPTO
CC at seqdata.uspto.gov/sequence.html?docId=20020197669.
XX
SQ Sequence 479 BP; 163 A; 107 C; 125 G; 84 T; 0 other;
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGAACCCCTGATGAAGAGCCCTAAAGAGAGAAACACCCACTAAAGTTCGGAA 378
Db 103 GTCCAGAGAGCCCACTAATGGAGAGCCCCAAAGAGAGAACAGCAGCTGAATCGGGA 162
QY 379 TCTCACCTGATCAGAGAGAGAGATGATCAGGGTCCAGTGCAGTATTC 428
Db 163 TCTCACCTGCGCAGCAGACAGAGAGATCAGGATACAGTGCAGTATTC 222
QY 429 -----AGTGCTGCTGAGCGGATCTCCAGGAGTATGTCAGACAAAGACTGGGG 482
Db 223 GAAGGGAATGCGGACATGGAGGTGATCTCAAGAGCTGATCAGTCAACACCGGG 282
QY 483 ATGGATGTAAGGT-GGTACTGATGTCAGGGGAGATTTCTACCAAAAGCAGACATTT 541
Db 283 ATAAATCTGGATTGGGTTCCGGGCTCAAGGTGAAGATAAATACCTAAAGAGGAACTGT 342
QY 542 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 601
Db 343 AAAATCCAGAGAGCGGTGAAGGGAATCAGAGGTTTAAAGGAAGTGAAGTGAACAC 402
QY 602 ACAAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655
Db 403 GCAAGCTGGTTTATATTAGGATATTTGACTTAAACTATCTCAATAAAGTTTG 456
Query Match 26.9%; Score 180; DB 25; Length 479;
Best Local Similarity 74.0%; Pred. No. 2.8e-44;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;

QY	379	TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTG	----- 428
Db	163	TCCTACACCTGGCAGCAGACAGAGAGAGATCAGGATCAGCTGAGATCCAGTCTGG	222
QY	429	-----AAGTGCCTGACTGAGAGCCGATCTCAGGAGCTATGTCAGACAAAGACTGGGG	482
Db	223	GAAAGGAAATCGCGACATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGG	282
QY	483	ATGGATGTGAAGGT-GGTACTGATGTCAGGGGAGAGATTCTACAAAAGCAGAGCACTTT	541
Db	283	ATAAATCTGGATTGGGCTCCGGGCTCAAGTGAAGTAACTAAAGAGGAACACTGT	342
QY	542	AAATGCCAGAGCAGGTGAAGGAAATCAAGGTTTAAAGGAGATAGCTGAACCAAC	601
Db	343	AAATGCCAGAGCAGGTGAAGGAAATCAAGGTTTAAATGAAGCAAGCTGAACCAAC	402
QY	602	ACAACTGTTTATATATAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG	655
Db	403	GCAAGCTGGTTTATATAGATATTTGACTTTAACTATCTCAATAAGTTTGG	456
RESULT 41			
ID	ABX99343		
XX	ABX99343 standard; cDNA; 479 BP.		
AC	ABX99343;		
DT	22-MAY-2003 (first entry)		
DE	Lung cancer therapy and diagnosis associated cDNA #70.		
KW	Lung cancer; cytostatic; vaccine; gene therapy; cancer;		
OS	gene; ss.		
XX	Homo sapiens.		
XX	US2002172952-A1.		
XX	21-NOV-2002.		
XX	10-JUL-2001; 2001US-0902941.		
XX	30-JUN-1999; 99US-0346492.		
PR	15-OCT-1999; 99US-0419356.		
PR	17-DEC-1999; 99US-0466867.		
PR	30-DEC-1999; 99US-0476300.		
PR	06-MAR-2000; 2000US-0519642.		
PR	22-MAR-2000; 2000US-0533077.		
PR	10-APR-2000; 2000US-0546259.		
PR	27-APR-2000; 2000US-0560406.		
PR	05-JUN-2000; 2000US-0589184.		
PR	11-JUL-2000; 2000US-0614124.		
PR	29-AUG-2000; 2000US-0651563.		
PR	08-SEP-2000; 2000US-0658824.		
PR	26-SEP-2000; 2000US-0671325.		
PR	06-OCT-2000; 2000US-0677419.		
PR	30-OCT-2000; 2000US-0702705.		
PR	13-DEC-2000; 2000US-0736457.		
PR	03-MAY-2001; 2001US-0849626.		
XX	(CORI-) CORIXA CORP.		
XX	Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW, Durham M;		
PI	Carter D, Fanger GR, Vedvick TS, Bangur CS, McNabb A;		
XX	WPI; 2003-328427/31.		
XX	New polynucleotide, useful for preparing a composition for treating or		
PT	inhibiting development of cancer, e.g. lung cancer -		
XX	Example 1; SEQ ID NO 70; 82pp; English.		
PS			
XX			

CC	The invention describes an isolated polynucleotide comprising one of 32		
CC	sequences, complement or degenerate variants of them. The polynucleotide		
CC	is useful for preparing a composition e.g. a vaccine or for gene therapy,		
CC	for treating or inhibiting development of cancer, e.g. lung cancer.		
CC	This sequence represents a polynucleotide associated with the		
CC	compositions and methods for the therapy and diagnosis of lung cancer.		
XX			
SEQ	Sequence 479 BP; 163 A; 107 C; 125 G; 84 T; 0 other;		
Query Match		26.9%; Score 180; DB 25; Length 479;	
Best Local Similarity		74.0%; Pred. No. 2.8e-44;	
Matches 262; Conservative		0; Mismatches 75; Indels 17; Gaps 2;	
QY	319	GTCCGAGAACCCACTGATGAAGAGCCTTAAGAGAGAGAAACCCACTAAAGTCGAA 378	
Db	103	GTCCGAGAGCCAGTAATGGAGAGCCCAAAAGAGAACCCAGCAGCTGAAGTCGGGA 162	
QY	379	TCCTACACCTGATCAGAAAGAGAGAGATGATCAGGGTGCAGCTCAGATTG----- 428	
Db	163	TCTACACCTGGCAGCAGACAGAGAGATCAGGATACAGTCCAGTCCCGAGTGTGG 222	
QY	429	-----AAGTGCCTGACCTCGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGG 482	
Db	223	GAAAGGAAATCGCGACATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGGG 282	
QY	483	ATGGATGTGAAGGT-GGTACTGATGTCAGGGGAGAGATTCTACAAAAGCAGAGCACTTT 541	
Db	283	ATAAATCTGGATTGGGCTCCGGGCTCAAGTGAAGTAACTAAAGAGGAACACTGT 342	
QY	542	AAATGCCAGAGCAGGTGAAGGAAATCAAGGTTTAAAGGAGATAGCTGAACCAAC 601	
Db	343	AAATGCCAGAGCAGGTGAAGGAAATCAAGGTTTAAATGAAGCAAGCTGAACCAAC 402	
QY	602	ACAACTGTTTATATATAGATATTTTACTTTAAAGAGTCTTAAATAATTTTG 655	
Db	403	GCAAGCTGGTTTATATAGATATTTGACTTTAACTATCTCAATAAGTTTGG 456	
RESULT 42			
ID	AAS60826		
XX	AAS60826 standard; cDNA; 1245 BP.		
AC	AAS60826;		
DT	29-JAN-2002 (first entry)		
DE	Human cancer agent-sensitive marker #328.		
XX	Human; cancer cell marker; TAXOL; cytostatic; tumour; carcinoma;		
KW	squamous cell carcinoma; sarcoma; fibrosarcoma; leukaemia;		
KW	lymphocytic leukaemia; lymphoma; plasmocytoma; reticulum cell sarcoma;		
XX	Hodgkin's disease; glioma; ss.		
OS	Homo sapiens.		
XX	WO200179556-A2.		
XX	25-OCT-2001.		
XX	13-APR-2001; 2001WO-US12132.		
XX	14-APR-2000; 2000US-197538P.		
XX	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.		
PI	Lillie J, Brown JL, Bolt A, Van Huffel C;		
XX	WPI; 2001-602933/68.		
XX	Novel nucleic acid, used as a marker to determine the effectiveness of		
PT	using TAXOL to treat cancer cell growth in individuals -		
XX	Claim 1; Page 340; 527pp; English.		
PS			
XX			

XX The invention relates to 1046 novel nucleic acids which are used as
CC markers for determining the sensitivity of a cancer cell to the
CC anticancer agent TAXOL. Cancer cells can be treated with TAXOL when
CC they are shown to express one of the 242 sensitivity markers or the
CC cells are shown not to express one of the 804 resistance markers.
CC The methods can be used to determine the effectiveness of TAXOL
CC in the treatment of cancer cell growth in an individual. The markers
CC can be used as targets in developing anti-cancer agents such as
CC chemotherapeutic compounds. The markers can also be used as targets in
CC developing treatments for cancer, particularly those cancers which
CC display resistance to agents and exhibit expression of the markers. The
CC anticancer agents developed by the novel method can be used to treat
CC cancer. Probes based on the markers can be used to detect transcripts or
CC genomic sequences corresponding to the markers, in the identification of
CC cells or tissues which mis-express the protein. Cancers which may
CC be targeted include carcinoma (e.g. squamous cell carcinoma),
CC sarcoma (e.g. fibrosarcoma) leukaemia (e.g. lymphocytic leukaemia),
CC lymphoma, plasmacytoma, reticulum cell sarcoma, Hodgkin's disease and
CC tumours (e.g. glioma). The present sequence is one of the 1046
CC novel cancer cell markers.
XX
SQ Sequence 1245 BP; 382 A; 308 C; 390 G; 153 T; 12 other;
Query Match 26.8%; Score 179.6; DB 22; Length 1245;
Best Local Similarity 67.1%; Pred. No. 6.1e-44;
Matches 286; Conservative 0; Mismatches 134; Indels 6; Gaps 2;
QY 217 GTGAATATGAGTTCGGAGGAAGATCAACATATAGGCTAGGCCAAGAAAGATTACA 276
Db 255 GTGAATATGAGTTCGGAGGAAGATCAACATATAGGCTAGGCCAAGAGCTACGTAGA 314
QY 277 GCCTCTGAGCTGATGGGGCTATGCTTATCGGCTCCCTTTGTCAGGAACCCACTGA 336
Db 315 GCCTCTGAAATGATGGGGCTATGCGGCC---CGAGCAGTTCAGTGATGAAGTGAACC 371
QY 337 TGAAGAGCTTAAGAGAAACACCCCACTAAAGTCGGAATCCTACACCTGATCAGAA 396
Db 372 AGCAACACTGAGAAAGGAGACCACTCAAGTCAGGATCTGAGCTGCTCAGGA 431
QY 397 GAGAGAAGATGATCAGGTCAGCTGAGATTCAAGTGCCTGACCTGGAGCGGATCTCCA 456
Db 432 GGGAGAGGATG---AGGAGCATCTGTCAGGTCAAGGGCCGAGCTCAAGTCATAGCCA 488
QY 457 GGAGCTATCTCAGACAAAGCTGGGATGATGATGATGATGATGATGATGATGATGATG 516
Db 489 GGAACAGGTCACCCACAGCTGGGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 548
QY 517 GATTCTACAAAGCAGACACTTTAAATGCGAAGCAGGTGAAGGGAATCAAGGT 576
Db 549 GGACCCGCCAATCCAGAGGAGGTGAACACGCTGAAGAGGTGAAGAGCAATCAAGTG 608
QY 577 TTAAGGAGATAGCTGAACACACAACTGTTTTATATTAGATATTTTACTTTAA 636
Db 609 TTAAGAGAGGCACGTTGAAATGATGACGCTCTCTATGTTGAAATTTGTCATTAA 668
QY 637 GAGTCT 642
Db 669 AATTCT 674
RESULT 43
AAF68852
ID AAF68852 standard; cDNA; 461 BP.
XX
AC AAF68852;
XX
DT 12-APR-2001 (first entry)
XX
DE Human lung tumour protein related nucleotide sequence SEQ ID NO:792.
XX
KW Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
KW lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;

KW cytostatic; antisense inhibition; ss.
XX Homo sapiens.
OS
XX WO200100828-A2.
PN
XX
PD
XX 04-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18061.
XX
XX 30-JUN-1999; 99US-0346492.
PR 15-OCT-1999; 99US-0419356.
PR 17-DEC-1999; 99US-0468667.
PR 30-DEC-1999; 99US-0476300.
PR 06-MAR-2000; 2000US-0519642.
PR 22-MAR-2000; 2000US-0533077.
PR 10-APR-2000; 2000US-0546259.
PR 27-APR-2000; 2000US-0560406.
PR 05-JUN-2000; 2000US-0589184.
XX
XX (CORI-) CORIXA CORP.
XX
XX Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
PI Retter MW, Mannion J;
PI
XX WPI; 2001-071488/08.
XX
DR Lung tumor-associated proteins and the nucleic acids that encode them,
PT useful for preventing, diagnosing and treating lung cancer -
PT
XX
XX
PS Claim 4; Page 413; 436pp; English.
XX
CC The present invention describes immunogenic portions of lung tumour-
CC associated proteins (I) and the nucleic acids (NAs) that encode them.
CC (I) have cytosolic activity and can be used in gene therapy, antisense
CC inhibition and in vaccines. The NAs and the lung tumour-associated
CC proteins they encode may be used in the prevention, treatment and
CC diagnosis of diseases associated with their inappropriate expression,
CC especially lung cancers. For example, the NAs may be administered to
CC treat diseases by rectifying mutations or deletions in a patient's genome
CC that affect the activity of the protein by expressing inactive proteins
CC or to supplement the patient's own production of (I). Additionally, the
CC NAs may be used to produce the lung-tumour associated protein, according
CC to standard recombinant DNA methodology. Conversely, antisense NA
CC molecules may be administered to down regulate protein expression by
CC binding with the cells own genes and preventing their expression. The NA
CC and complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar NA sequences in
CC samples, and hence which patients may be in need of treatment for lung
CC cancer. The (I) may be used as antigens in the production of antibodies
CC and in assays to identify modulators (agonists and antagonists) of the
CC expression and activity of the protein. AAF68083 to AAF68878 and
CC AAB76848 to AAB76878 represent human lung tumour protein related
CC nucleotide and protein sequences which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 461 BP; 150 A; 104 C; 123 G; 84 T; 0 other;
Query Match 26.6%; Score 178; DB 22; Length 461;
Best Local Similarity 73.9%; Pred. No. 1.1e-43;
Matches 260; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
QY 319 GTCCAGGACCCACTGATCAAGAGCCTAAAGAGAGAGAGATGATGAGGTGAGCTGAGATTC 378
Db 101 GTCCAGGAGCCCATGTAATGGAGAGCCCAAAAGAGAGAGAGATGAGTGAAGTGGGA 160
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATGAGGTGAGCTGAGATTCAGGTGCTGG 428
Db 161 TCCTACACCTGGGAGAGAGAGATGATGAGGTGAGCTGAGATTCAGGTGCTGG 220
QY 429 -----AAGTGCCTGACCTGGAGCCGATCTCCAGGAGCTATGTCACACAAAGACTGGGG 482
Db 221 GAAGGGAATGCGCGCATGGAAGGTGATTTGCAAGAGCTGCTCATGCTCAACACCGGG 280

QY 483 ATGATGTGAAGGT-GGTACTGATGTCAAGGGGAAGATTCTACCAAAAGCAGACACTTT 541
 DB |||||
 DB 281 ATAAATCTGGATTGGGTTCCGGCTCAAGGTGAAGATAATACCTAAAGAGGAACACTGT 340
 QY 542 AAAATGCCAAGCAGGTGAAGGGAATCAACAGTTTAAAGGAAGATAGCTGAACAAC 601
 DB |||||
 DB 341 AAAATGCCAAGCAGGTGAAGGGAATCAACAGTTTAAAGGAAGATAGCTGAACAAC 400
 QY 602 ACAAACTGTTTTATATAGATATTTTAAAGAGTCTTAAATAATTTT 653
 DB |||||
 DB 401 GCAAGCTGGTTTTATATAGATATTTTAAAGAGTCTTAAATAATTTT 452

RESULT 44

ABK38763

ID ABK38763 standard; cdNA; 461 BP.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

Query Match 26.6%; Score 178; DB 24; Length 461;
 Best Local Similarity 73.9%; Pred. No. 1.1e-43;
 Matches 260; Conservative 0; Mismatches 75; Indels 17; Gaps 2;
 QY 319 GTCCAGCAACCCCTGATGAAGAGCCTTAAGAGAGAGAAACCCCTACTAAAGTCGGAA 378
 DB |||||
 DB 101 GTCCAGCAACCCCTGATGAAGAGCCTTAAGAGAGAGAAACCCCTACTAAAGTCGGAA 160
 QY 379 TCCTACACCTGATCAGAAAGAGAGATGATCAGGGTGCAGTGATTC----- 428
 DB |||||
 DB 161 TCCTACACCTGCGCAGCAGACAGAAAGATCAGGATACAGTATCCAGGTGCTGG 220
 QY 429 -----AAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGCAAGAGCTGGG 482
 DB |||||
 DB 221 GAAGGGAATGCCGACATGGAAGTGAITCTGCAAGAGCTGCATCAGTCNAACCCGGG 280
 QY 483 ATGATGTGAAGGT-GGTACTGATGTCAAGGGGAAGATTCTACCAAAAGCAGACACTTT 541
 DB |||||
 DB 281 ATAAATCTGGATTGGGTTCCGGCTCAAGGTGAAGATAATACCTAAAGAGGAACACTGT 340
 QY 542 AAAATGCCAAGCAGGTGAAGGGAATCAACAGTTTAAAGGAAGATAGCTGAACAAC 601
 DB |||||
 DB 341 AAAATGCCAAGCAGGTGAAGGGAATCAACAGTTTAAAGGAAGATAGCTGAACAAC 400
 QY 602 ACAAACTGTTTTATATAGATATTTTAAAGAGTCTTAAATAATTTT 653
 DB |||||
 DB 401 GCAAGCTGGTTTTATATAGATATTTTAAAGAGTCTTAAATAATTTT 452

RESULT 45

ACAL1092

ID ACAL1092 standard; cdNA; 461 BP.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

(CORI-) CORIXA CORP.

Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;

Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS;

McNabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;

WPI; 2002-164634/21.

P-PSDB; AAU85524.

Novel polynucleotide encoding a lung tumour polypeptide useful for

stimulating and/or expanding T cells specific for a tumour protein

-

Example 1; SEQ ID No 792; 223pp; English.

The invention describes an isolated polynucleotide and polypeptide

useful for stimulating and/or expanding T cells specific for a tumour

protein for determining the presence of a cancer in a patient. A

composition containing the polynucleotide and/or polypeptide is useful

for treating a lung cancer in a patient. The polynucleotide is

removing tumour cells from a biological sample. The polynucleotide is

also useful as probe or primer to detect the level of mRNA encoding a

tumour protein. This sequence encodes a lung tumour associated protein

or protein fragment, described in the method of the invention.

Note: The sequence data for this patent did not form part of the

printed specification, but was obtained in electronic format directly

from WIPO at

ftp.wipo.int/pub/published_pct_sequences.

Sequence 461 BP; 150 A; 104 C; 123 G; 84 T; 0 other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Novel lung cancer polynucleotide encoding lung cancer protein, useful
 for detecting the presence of lung cancer in a patient, and in
 pharmaceutical compositions, e.g. vaccines, for treating lung cancer
 Example 1; Page -; 72pp; English.

XX

Search completed: January 18, 2004, 01:58:21

Job time : 130.524 secs

CC The invention relates to a polynucleotide encoding a lung tumour protein,
CC comprising a sequence selected from any of the 14 sequences
CC mentioned in the specification, or a sequence (S2) mentioned in
CC specification, complement of S1, sequences consisting of at least 20
CC contiguous residues of S1, sequences that hybridise to S1, sequences
CC having 75%, preferably 90%, identity to S1, or degenerate variants of
CC S1. Also included are an isolated polypeptide (comprising a sequence (S3)
CC selected from any one of the 4 amino acid sequences mentioned in the
CC specification, a sequence encoded by the polynucleotide, or sequences
CC having at least 70%, preferably 90%, identity to a sequence encoded by
CC the polynucleotide), an expression vector comprising the polynucleotide
CC operably linked to an expression control sequence, a host cell
CC transformed or transfected with the vector, an isolated antibody (or its
CC antigen-binding fragment) that specifically binds to the polypeptide,
CC detecting the presence of a cancer in a patient, a fusion protein
CC comprising the polypeptide, an oligonucleotide that hybridises to
CC S1 under moderately stringent conditions, stimulating and/or expanding T
CC cells specific for a tumour protein (comprising contacting T cells with
CC the polynucleotide, protein or antigen-presenting cells, under conditions
CC and for a time sufficient to permit the stimulation and/or expansion of T
CC cells) and inhibiting the development of a cancer in a patient (by
CC incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with the
CC polynucleotide, protein or antigen presenting cells that express the
CC polynucleotide, such that T cells proliferate, administering to the
CC patient an effective amount of the proliferated T cells, and thus
CC inhibiting the development of a cancer in the patient. The
CC polynucleotide, protein and cells are useful in a composition for
CC stimulating an immune response in a patient, and for treating a cancer in
CC a patient (particularly lung cancer). The oligonucleotide is useful for
CC determining the presence of a cancer in a patient. The protein and
CC oligonucleotides are useful in pharmaceutical compositions, e.g.
CC vaccines. The polynucleotide is also useful as a probe or primer for
CC nucleic acid hybridisation, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of tumour polypeptides and
CC proteins in tumour cells. An amplified portion of the polynucleotide is
CC useful for isolating a full-length gene from a suitable library.
CC The present sequence is a cDNA (full length, extended or partial)
CC isolated from a library derived from lung tumour/cancer cells.
CC Note: the sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from the USPTO
CC at seqdata.uspto.gov/sequence.html?DocId=20020197669.

XX
SQ Sequence 461 BP; 150 A; 104 C; 123 G; 84 T; 0 other;

Query Match		26.6%;	Score 178;	DB 25;	Length 461;
Best Local Similarity		73.9%;	Pred. No. 1.1e-43;		
Matches 260;		Conservative 0;	Mismatches 75;	Indels 17;	Gaps 2;
QY	319	GTCCAGGAAACCCACTGTATGAAGAGCCTAAGAGAGAGAAACCCACCTAAAGTGGAA	378		
Db	101	GTCCAGGAGCCAGTATGTGGAGAGCCCAAAAGAGAACCCAGCAGCTGAAAGTCGGGA	160		
QY	379	TCTACACCTGATCAGAGAGAGAGATGATCAGGTGCGCTGAGATTC-----	428		
Db	161	TCCTACCTGGCAGCAGACAGAGAGATCAGATCAGCTGAGATCCCGAGTGTGG	220		
QY	429	-----AAGTGCCTGACCTGGAGCCGATCTCCAGGAGCTATGTCTAGACAAAGACTGGGG	482		
Db	221	GAAGGGAATGCGCAGATGGAAGTGTATCTCAGAGCTGCATCAGTCAACACCGGGG	280		
QY	483	ATCGATGTGAAGGT-GGTACTGATGTCAAGGGGAGAGATTCTACCAAGACAGCAGCTTT	541		
Db	281	ATAAATCTGGATTGGGTTCCGGCGTCAAGGTGAAGATTAATACCTTAAGAGGAACACTGT	340		
QY	542	AAATGCGCAGAGCAGGTGAAGGGAATCAGAGTTTAAAGGAAGATAGCTGAAACAC	601		
Db	341	AAATGCCAGAGCAGGTGAAGGCAACCAAGTTTAATGAAGAGAGCTGAAACAC	400		
QY	602	ACAAACGTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAATAATTTT	653		
Db	401	GCAAGCTGGTTTTATATTAGATATTTGACTTAAACTATCTCAATAAAGTTT	452		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2004, 01:12:25 ; Search time 32.4394 Seconds
(without alignments)
9102.683 Million cell updates/sec

Title: US-10-051-835-1

Perfect score: 669

Sequence: 1 cacaacgcagcagcagcagtt.....tttttggcgtgcgtgcattc 669

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139556

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*

5: /cgn2_6/ptodata/2/ina/PCRTUS COMB.seq.*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	317	47.4	659	4	US-09-620-312D-634
2	204	30.5	399	4	US-09-702-705-69
3	204	30.5	399	4	US-09-736-457-69
4	204	30.5	457	4	US-09-702-705-790
5	204	30.5	457	4	US-09-736-457-790
6	204	30.5	781	4	US-09-702-705-808
7	204	30.5	781	4	US-09-736-457-808
8	182.8	27.3	528	4	US-09-163-748C-1
9	181.2	27.1	538	1	US-08-370-648-14
10	181.2	27.1	538	2	US-08-531-662B-14
11	181.2	27.1	538	3	US-08-602-039-14
12	181.2	27.1	538	3	US-08-602-039-14
13	180	26.9	479	4	US-09-702-705-70
14	180	26.9	479	4	US-09-736-457-70
15	178	26.6	461	4	US-09-702-705-792
16	178	26.6	461	4	US-09-736-457-792
17	169.8	25.4	532	1	US-08-370-648-17
18	169.8	25.4	532	2	US-08-531-662B-17
19	169.8	25.4	532	3	US-08-602-039-17
20	169.8	25.4	532	3	US-08-602-039-17
21	169.8	25.4	539	1	US-08-370-648-18
22	169.8	25.4	539	2	US-08-531-662B-18
23	169.8	25.4	539	3	US-08-602-039-18
24	169.8	25.4	539	3	US-08-602-039-18
25	169.8	25.4	540	1	US-08-370-648-16
26	169.8	25.4	540	2	US-08-531-662B-16
27	169.8	25.4	540	3	US-08-602-039-16

28	169.8	25.4	540	3	US-08-602-039-16
29	168.2	25.1	526	4	US-09-163-748C-2
30	162.2	24.2	560	1	US-08-370-648-15
31	162.2	24.2	560	2	US-08-531-662B-15
32	162.2	24.2	560	3	US-08-602-039-15
33	162.2	24.2	560	3	US-08-602-039-15
34	155.2	23.2	288	4	US-09-702-705-1827
35	155.2	23.2	288	4	US-09-736-457-1827
36	148.6	22.2	656	4	US-09-328-475C-263
37	144.2	21.6	646	1	US-08-370-648-1
38	144.2	21.6	646	2	US-08-531-662B-1
39	144.2	21.6	646	3	US-08-602-039-1
40	144.2	21.6	646	3	US-08-602-039-1
41	144.2	21.6	646	3	US-08-602-039-1
42	142.8	21.0	1024	4	US-09-328-475C-78
43	140.8	21.0	1024	4	US-09-328-475C-77
44	99.2	14.8	515	4	US-08-439-313-472
45	99.2	14.8	515	4	US-09-352-616A-472
46	97	14.5	9531	4	US-09-163-748C-3
47	97	14.5	11461	3	US-08-602-039-1
48	78	11.7	1092	4	US-09-712-016-5
49	76.2	11.4	141	4	US-09-702-705-1828
50	76.2	11.4	141	4	US-09-736-457-1828
51	60.6	9.1	7218	1	US-08-232-463-14
52	42.8	6.4	111	4	US-09-702-705-1829
53	42.8	6.4	111	4	US-09-736-457-1829
54	39.6	5.9	289	3	US-09-007-005-17
55	39.6	5.9	289	3	US-09-244-796-17
56	35.4	5.3	966	2	US-08-766-738-2
57	35.4	5.3	966	2	US-09-262-610-2
58	34.6	5.2	246	4	US-09-407-956-3
59	34.6	5.2	888	4	US-09-407-956-1
60	33.6	5.0	28473	4	US-08-961-527-83
61	33.4	5.0	139608	4	US-09-539-333D-1
62	33.4	5.0	1230025	4	US-09-198-452A-1
63	33.2	5.0	6678	4	US-09-268-347-33
64	33	4.9	822	3	US-08-896-164-47
65	33	4.9	1344	4	US-09-480-921B-5
66	33	4.9	4140	4	US-09-480-921B-19
67	33	4.9	7739	3	US-09-195-956-1
68	33	4.9	7745	1	US-08-299-675-1
69	33	4.9	7745	1	US-08-485-241-1
70	33	4.9	7745	2	US-08-874-162-1
71	33	4.9	61663	4	US-09-453-702B-62
72	32.8	4.9	325	1	US-08-650-275-8
73	32.8	4.9	325	3	US-09-181-318-8
74	32.8	4.9	374	1	US-08-341-538A-3
75	32.8	4.9	374	2	US-08-725-518-3
76	32.8	4.9	1113	1	US-08-341-538A-1
77	32.8	4.9	1113	2	US-08-725-518-1
78	32.8	4.9	1323	4	US-09-252-991A-5119
79	32.8	4.9	1422	4	US-09-252-991A-5119
80	32.8	4.9	1467	4	US-09-252-991A-5070
81	32.8	4.9	1664976	4	US-08-916-421B-1
82	32.6	4.9	5361	3	US-08-973-462-2
83	32.6	4.9	5361	3	US-08-973-462-1
84	32.2	4.8	277	4	US-09-313-294A-6369
85	32.2	4.8	319608	4	US-09-679-409-1
86	32	4.8	1551	1	US-08-485-241-4
87	32	4.8	1551	2	US-08-874-162-4
88	32	4.8	1919	1	US-08-485-241-2
89	32	4.8	1919	2	US-08-874-162-2
90	32	4.8	3334	4	US-09-996-243-288
91	31.8	4.8	394	1	US-08-650-275-12
92	31.8	4.8	394	3	US-09-181-318-12
93	31.8	4.8	2277	1	US-08-676-967-2
94	31.8	4.8	2277	1	US-08-676-974-2
95	31.8	4.8	2277	2	US-09-096-487-2
96	31.8	4.8	3592	3	US-08-714-918-63
97	31.8	4.8	3592	3	US-09-265-315-63
98	31.8	4.8	3592	3	US-09-265-315-63
99	31.8	4.8	3592	3	US-09-265-315-63
100	31.6	4.7	429	4	US-09-643-597-199

Sequence 16, Appli
Sequence 2, Appli
Sequence 15, Appli
Sequence 15, Appli
Sequence 15, Appli
Sequence 15, Appli
Sequence 1827, Ap
Sequence 1827, Ap
Sequence 263, App
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 78, Appli
Sequence 77, Appli
Sequence 472, App
Sequence 472, App
Sequence 3, Appli
Sequence 29, Appli
Sequence 5, Appli
Sequence 1828, Ap
Sequence 1828, Ap
Sequence 14, Appli
Sequence 1829, Ap
Sequence 1829, Ap
Sequence 17, Appli
Sequence 2, Appli
Sequence 3, Appli
Sequence 83, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 62, Appli
Sequence 8, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 5147, Ap
Sequence 5119, Ap
Sequence 5070, Ap
Sequence 1, Appli
Sequence 2, Appli
Sequence 1, Appli
Sequence 6369, Ap
Sequence 1, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 288, App
Sequence 12, Appli
Sequence 12, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 63, Appli
Sequence 63, Appli
Sequence 63, Appli
Sequence 199, App

ALIGNMENTS

RESULT 1
US-03-620-312D-634
; Sequence 634, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Dmanac, Radje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt FL_Genes Version 1.0
; SEQ ID NO 634
; LENGTH: 659
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (225)..(560)
US-09-620-312D-634

Query Match 47.4%; Score 317; DB 4; Length 659;
Best Local Similarity 77.4%; Pred. No. 1.6e-91;
Matches 442; Conservative 0; Mismatches 85; Indels 44; Gaps 3;
QY 86 CTTTGGCCACGTGTGACCGCTGGGAGCTGTGAGAGTGTGAGGGGACACCTTCCAGCG 145
DB 109 CTTTTCACCGTGTGATGCCCTCGAGTTGTGAGGGTGTGAGGGTGGCGTTCCTGCTG 168
QY 146 TCTGGACTTTCT 205
DB 169 TCTGACATTTTCT 210
QY 206 GAACCTGAATAGTGAATATAGTTGGGAGGAGGAGATCAACATATAGCCCTAGGCCAAGA 265
DB 211 ----TTGCACTGTGAATATGATTGGGAGGAGGAGATCAACATATAGGCCCTAGGCCAGG 266
QY 266 AGAAGTTACAGCCCT 325
DB 267 AGAAGTTACAGCCCT 305
QY 326 GAACCCACTGTATGAAGAGCTTAAGAAGAAACACCCACTTAAAGTCGGAATCCTACA 385
DB 306 GAGCCCGGTGTAGGAGGCTCAGCAAGAGGAAACCACTGAAGTCGGGATCCTGCA 365
QY 386 CTTGATCAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGTGCCTGACTCGAA 445

RESULT 2

US-09-702-705-69
; Sequence 69, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 69
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-69

Query Match 30.5%; Score 204; DB 4; Length 399;
Best Local Similarity 77.4%; Pred. No. 1.7e-55;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGAAACCCACTGTGTAAGAGCCTTAAAGAGAGAAACCCACCCTTAAAGTCGGA 378
DB 40 GTCCAGGAGCCAGTAATGGAGAGCCCAAAAGAAAGAACCCAGCAGCTGAAAGTCGGA 99
QY 379 TCCTACACCTGATCAGAAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGATTCGCTGA 438
DB 100 TCCTACACCTGGCAGCAGACAGAGAGATCAGATACAGCTGAGATCCCAAGTCGGA 159
QY 439 CTTGGAAGCCGATCTCCAGAGCTATGTGAGCAAGAGCTGGGGATGGATGTAGAGGT-G 497
DB 160 CATGGAAGGTGATCTGCAAGAGCTGATCAGTCAAAACACCCGGGATAAATCTGGATTGG 219
QY 498 GTACTGATGTCAAGGGGAAGATTCTACAAAAGCAGAGCAGCTTTAAATGCCAGAGCAG 557
DB 220 GTTCCGGCTCAAGGTGAAGATAATACCTTAAGAGGAACACTGTAAATGCCAGAGCAG 279
QY 558 GTGAAGGGAATCAGAGGTTTAAAGGAAGATAGCTGAAGCAACCAAACTGTTTTTATA 617
DB 280 GTGAAGAGCAACCAAGTTTAAATGAAGCAAGCTGAAGCAACCAAGCTGTTTTTATA 339
QY 618 TTAGATATTTTACTTTTAAAGAGTCTTTAATAAATTTT 653
DB 340 TTAGATATTGACTTAAACTATCTCAATAAAGTTTTT 375

RESULT 3

US-09-736-457-69
; Sequence 69, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 69
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-69

Query Match 30.5%; Score 204; DB 4; Length 399;
Best Local Similarity 77.4%; Pred. No. 1.7e-55;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

QY 319 GTCCAGGAAACCCACTGATGAAGAGCCTTAAGAGAGAACCCACCACCTAAAGTCCGAA 378
DB 40 GTCCAGGAGCCCAAGTAAATGGAGAGCCCAAAAGAGAACCCAGCAGCTGAAGTCCGGA 99

QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGTGCCTGA 438
DB 100 TCCTACACCTGGGAGAGAGAGATGATCAGGATCAGGATCAGATCCAGTGCCTGA 159

QY 439 CCTGGAAGCCGATCTCAGAGAGCTATGTCAGACAAAGACTGGGGATGGATGTAAGT-G 497
DB 160 CATGGAAGTGTATCTGCAAGAGCTGCATCAGTCAACACCCGGGATAAATCTCGATTTGG 219

QY 498 GTACTGATGTCAAGGGAAGATTCTACAAAGCAGAGCAGCTTTAAATGCCAGAGCAG 557
DB 220 GTTCGGGTCAGGTGAGATTAATACCTAAAGAGGAACTGTAAATGCCAGAGCAG 279

QY 558 GTGAAGGAAATCAAGTTTAAAGGAGATAAGCTGAAACAAACAACTGTTTTTATA 617
DB 280 GTGAAGAGCAACCAAGTTTAAATGAAGACAGCTGAAACAAAGCTGTTGTTTTATA 339

QY 618 TTAGTATTTTACTTTAAAGAGTCTTAATAATTTT 653
DB 340 TTAGTATTTGACTTAAACTATCTCAATAAAGTTTT 375

RESULT 4

US-09-702-705-790
; Sequence 790, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 790
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-790

Query Match 30.5%; Score 204; DB 4; Length 457;
Best Local Similarity 77.4%; Pred. No. 1.8e-55;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

QY 319 GTCCAGGAAACCCACTGATGAAGAGCCTTAAGAGAGAACCCACCCTAAAGTCCGAA 378
DB 120 GTCCAGGAGCCCAAGTAAATGGAGAGCCCAAAAGAGAACCCAGCAGCTGAAGTCCGGA 179

QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGTGCCTGA 438
DB 180 TCCTACACCTGGGAGAGAGAGATCAGGATCAGGATCAGATCCAGTGCCTGA 239

QY 439 CCTGGAAGCCGATCTCAGAGAGCTATGTCAGACAAAGACTGGGGATGGATGTAAGT-G 497
DB 240 CATGGAAGTGTATCTGCAAGAGCTGCATCAGTCAACACCCGGGATAAATCTGGATTTGG 299

QY 498 GTACTGATGTCAAGGGAAGATTCTACAAAGCAGAGCAGCTTTAAATGCCAGAGCAG 557
DB 300 GTTCGGGTCAGGTGAGATTAATACCTAAAGAGGAACTGTAAATGCCAGAGCAG 359

QY 558 GTGAAGGAAATCAAGTTTAAAGGAGATAAGCTGAAACAAACAACTGTTTTTATA 617
DB 360 GTGAAGAGCAACCAAGTTTAAATGAAGACAGCTGAAACAAAGCTGTTGTTTTATA 419

QY 618 TTAGTATTTTACTTTAAAGAGTCTTAATAATTTT 653
DB 420 TTAGTATTTGACTTAAACTATCTCAATAAAGTTTT 455

RESULT 5

US-09-736-457-790
; Sequence 790, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 790
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-790

Query Match 30.5%; Score 204; DB 4; Length 457;
Best Local Similarity 77.4%; Pred. No. 1.8e-55;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

QY 319 GTCCAGGAGCCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCTAAAGAGTGGGA 378
DB 120 GTCCAGGAGCCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCTAAAGAGTGGGA 179
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGAGTGGGTCAGCTGAGATTCAGTGCCTGA 438
DB 180 TCCTACACCTGATCAGAGAGAGAGATGATCAGAGTGGGTCAGCTGAGATTCAGTGCCTGA 239
QY 439 CCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGGATGTGAAGGT-G 497
DB 240 CATGGAGGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGATAAATCTGGATTGG 299
QY 498 GTACTGATGTCAGGAGAGATTTTACCAAGAGAGAGACTTTAAATGCCAGAGCAG 557
DB 300 GTTCGGCGTCAAGGTGAAGATAATACCTAAAGAGAGAGACTGTAAATGCCAGAGCAG 359
QY 558 GTGAAGGGAATACAGAGTGTAAAGAGAGATGATGAAACCAACCAAGCTGTTTTTATA 617
DB 360 GTGAAGAGCAACCAAGTTTAAATGAAGACAAGCTGAAACCAAGCTGTTTTTATA 419
QY 618 TTAGATATTTTACCTTAAAGAGCTCTTAATAAATTTT 653
DB 420 TTAGATATTTGACTTAAACTATCTCAATAAAGTTTTT 455

RESULT 5
US-09-702-705-808
; Sequence 808, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT FILING DATE: 2000-10-30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 808
; LENGTH: 781
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-808

Query Match 30.5%; Score 204; DB 4; Length 781;
Best Local Similarity 77.4%; Pred. No. 2.5e-55;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGAGCCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCTAAAGAGTGGGA 378
DB 422 GTCCAGGAGCCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCTAAAGAGTGGGA 481
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCAGCTGAGATTCAGTGCCTGA 438
DB 482 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCAGCTGAGATTCAGTGCCTGA 541
QY 439 CCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGGATGTGAAGGT-G 497
DB 542 CATGGAGGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGATAAATCTGGATTGG 601
QY 498 GTACTGATGTCAGGAGAGATTTTACCAAGAGAGAGACTTTAAATGCCAGAGCAG 557
DB 602 GTTCGGCGTCAAGGTGAAGATAATACCTAAAGAGAGAGACTGTAAATGCCAGAGCAG 661

QY 558 GTGAAGGGAATACAGAGTGTAAAGAGAGATGATGAAACCAACCAAGCTGTTTTTATA 617
DB 662 GTGAAGGGAATACAGAGTGTAAAGAGAGATGATGAAACCAACCAAGCTGTTTTTATA 721
QY 618 TTAGATATTTTACCTTAAAGAGCTCTTAATAAATTTT 653
DB 722 TTAGATATTTGACTTAAACTATCTCAATAAAGTTTTT 757

RESULT 7
US-09-736-457-808
; Sequence 808, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT FILING DATE: 2000-12-13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 808
; LENGTH: 781
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-808

Query Match 30.5%; Score 204; DB 4; Length 781;
Best Local Similarity 77.4%; Pred. No. 2.5e-55;
Matches 260; Conservative 0; Mismatches 75; Indels 1; Gaps 1;
QY 319 GTCCAGGAGCCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCTAAAGAGTGGGA 378
DB 422 GTCCAGGAGCCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCTAAAGAGTGGGA 481
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCAGCTGAGATTCAGTGCCTGA 438
DB 482 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTCAGCTGAGATTCAGTGCCTGA 541
QY 439 CCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGGATGTGAAGGT-G 497
DB 542 CATGGAGGTGATCTGCAAGAGCTGCATCAGTCAACACCGGGATAAATCTGGATTGG 601
QY 498 GTACTGATGTCAGGAGAGATTTTACCAAGAGAGAGACTTTAAATGCCAGAGCAG 557
DB 602 GTTCGGCGTCAAGGTGAAGATAATACCTAAAGAGAGAGACTGTAAATGCCAGAGCAG 661
QY 558 GTGAAGGGAATACAGAGTGTAAAGAGAGATGATGAAACCAACCAAGCTGTTTTTATA 617
DB 662 GTGAAGGGAATACAGAGTGTAAAGAGAGATGATGAAACCAACCAAGCTGTTTTTATA 721
QY 618 TTAGATATTTTACCTTAAAGAGCTCTTAATAAATTTT 653
DB 722 TTAGATATTTGACTTAAACTATCTCAATAAAGTTTTT 757

RESULT 8
US-09-163-748C-1
; Sequence 1, Application US/09163748C
; Patent No. 6509172
; GENERAL INFORMATION:
; APPLICANT: DeBacker, Oliver
; APPLICANT: Van den Eynde, Benoit

```
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Encode Member Of The Gage
; FILE REFERENCE: LUD 5558
; CURRENT APPLICATION NUMBER: US/09/163,748C
; CURRENT FILING DATE: 1998-09-30
; NUMBER OF SEQ ID NOS: 27
; SEQ ID NO 1
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-163-748C-1

Query Match      27.3%; Score 182.8; DB 4; Length 528;
Best Local Similarity 67.6%; Pred. No. 1.2e-48;
Matches 289; Conservative 0; Mismatches 132; Indels 6; Gaps 2;

QY 217 GTGAATATGAGTTGGCGAGAGATCAACATATAGCCCTAGCCCAAGAGAGTTTACA 276
Db 66 GTGAATATGAGTTGGCGAGAGATCGACCTATCGGCTAGACCAGCGCTACGTAGA 125

QY 277 GCCTCTGAGCTGATTGGGCTATGTTTACTTGCTCCCTTTGTCCTCCAGAACCCACTGA 336
Db 126 GCCTCTGAGATGATTGGGCTATGCGGCC---CGAGCAGTTTCAGTGAAGTGAACC 182

QY 337 TGAAGAGCCTAAAGAGAGAGAACCCACCCTAAAGTCGGAATCCTACACCTGATCAGAA 396
Db 183 AGCAACACCTGAGAGAGAGAGAACCGCACTCAAGTCAGGATCTCTGAGCTGCTCAGA 242

QY 397 GAGAGAGAGATGATCAGGGTCAGCTGAGATTCAGTTCAGTCCCTGAGAGCCGATCTCA 456
Db 243 GGGAGAGAGATG---AGGAGCAGCTGCGAGTCAAGGCGCGAAGCCTGAACTGATAGCA 299

QY 457 GGAGCTATGTCAGACAAAGACTGGGATGGATGATGAGGTGGTACTGATGTCAGGGGAA 516
Db 300 GGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGGTCCCTGATGGCGAGAGAT 359

QY 517 GATTCTACAAAAGCAGAGACACTTTAAATGCCAGAGCAGGTGAAGGGAATCAGAGT 576
Db 360 GGACCCGCCAAATCCAGAGAGGTGAAACGCCCTGAAAGAGTGAAGCAATCAGAGT 419

QY 577 TTAAGAGAGATAGCTGAACACACAACTGTTTATATATAGATATTTTACTTAA 636
Db 420 TTAAGAGAGACAGCTTGAATGATGACAGGCTGCTCTATGTGGAAATTTGTTCAATA 479

QY 637 GAGTCT 642
Db 480 AATTCT 485

RESULT 9
US-08-370-648-14
; Sequence 14, Application US/08370648
; Patent No. 5648226
; GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry;
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
```

```
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,648
; FILING DATE: 10-JANUARY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-JULY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5648226man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5323.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 538 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-370-648-14

Query Match      27.1%; Score 181.2; DB 1; Length 538;
Best Local Similarity 67.4%; Pred. No. 4e-48;
Matches 287; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

QY 217 GTGAATATGAGTTGGCGAGAGATCAACATATAGCCCTAGCCCAAGAGAGTTTACA 276
Db 77 GTGAATATGAGTTGGCGAGAGATCGACCTATCGGCTAGACCAGCGCTACGTAGA 136

QY 277 GCCTCTGAGCTGATTGGGCTATGCTTACTTGCTCCCTTTGTCCTCCAGGAACCCACTGA 336
Db 137 GCCTCTGAGATGATTGGGCTATGCGGCC---CGAGCAGTTTCAGTGAAGTGAACC 193

QY 337 TGAAGAGCCTAAAGAGAGAGAACCCACCCTAAAGTCGGAATCCTACCTGATCAGAA 396
Db 194 AGCAACACCTGAGAGAGAGAGAACCCAGCACTCAACAGCTCAGGATCCTGCGCTGCTCAGA 253

QY 397 GAGAGAGATGATCAGGGTCAGCTGAGATTCAGATTCAGTCCCTGAGAGCCGATCTCCA 456
Db 254 GGGAGAGAGATG---AGGAGCAGTCTGAGGTCAAGGCGCGAGCCTGAGCTCATAGCCA 310

QY 457 GGAGCTATGTCAGACAAAGACTGGGATGGATGTAAGGTGGTACTGATGTCAGAGGAA 516
Db 311 GGAACAGGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGCTCTGATGGCGAGAGAT 370

QY 517 GATTCTACAAAAGCAGAGCACTTTAAATGCCAGAGCAGGTGAAGGGAATCAGAGT 576
Db 371 GGACCCGCCAAATCCAGAGAGGTGAAACCGCTGAAAGAGGTGAAAGCAATCAGAGT 430

QY 577 TTAAGAGAGATAGCTGAAACACACAACTGTTTTATATAGATATTTTACTTTAA 636
Db 431 TTAAGAGAGACAGCTTGAATGATGACAGGCTGCTCTATGTTGGAATTTGTTCAATA 490

QY 637 GAGTCT 642
Db 491 AATTCT 496

RESULT 10
US-08-531-662B-14
; Sequence 14, Application US/08531662B
; Patent No. 5858689
; GENERAL INFORMATION:
; APPLICANT: Van der Bruggen, Pierre;
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated, Truncated Nucleic
; TITLE OF INVENTION: Acid Molecules Which Code For GAGE Tumor
```


QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTGCGAGTGCAGATTC----- 428
Db 163 TCCTACACCTGGGAGAGAGAGAGATGATCAGGTGCGAGTGCAGATTC----- 222
QY 429 -----AAGTGCCTGACCTGGAAGCGGATCTCCAGGAGCTATGTTCAGACAAAGAGACTGGG 482
Db 223 GAAGGGAATGCGGACATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGG 282
QY 483 ATGATGTGAAGGT-GGTACTGATGTCAAGGGGAAGTCTTACCAGAGAGAGACTTT 541
Db 283 ATAAATCTGGATTTGGGTTCCGGCTCAAGGTGAAGATAATACCTAAAGAGGAACTGT 342
QY 542 AAAATGCCAGAGAGAGGGAATCAAGGTCTTAAAGGAAGATGAAGCTGAAACAAAC 601
Db 343 AAAATGCCAGAGAGAGGGAATCAAGGTCTTAAAGGAAGATGAAGCTGAAACAAAC 402
QY 602 ACAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAATAATTTT 655
Db 403 GCAAGCTGGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAATAATTTT 456

RESULT 14
US-09-736-457-70
; Sequence 70, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 70
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-70

Query Match 26.9%; Score 180; DB 4; Length 479;
Best Local Similarity 74.0%; Pred. No. 9,1e-48;
Matches 262; Conservative 0; Mismatches 75; Indels 17; Gaps 2;

QY 319 GTCCAGGACCCACTGATGAAGAGCCTTAAGAGAGAGAAACCCACTTAAAGTCGAA 378
Db 103 GTCCAGGAGCCAGTAAATGGAGAGCCCAAAAGAGAAACAGCAGCTGAAGTCGGA 162
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTGCGAGTGCAGATTC----- 428
Db 163 TCCTACACCTGGGAGAGAGAGATGATCAGGTGCGAGTGCAGATTC----- 222
QY 429 -----AAGTGCCTGACCTGGAAGCGGATCTCCAGGAGCTATGTTCAGACAAAGACTGGG 482
Db 223 GAAGGGAATGCGGACATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGG 282
QY 483 ATGATGTGAAGGT-GGTACTGATGTCAAGGGGAAGTCTTACCAGAGAGAGACTTT 541
Db 283 ATAAATCTGGATTTGGGTTCCGGCTCAAGGTGAAGATAATACCTAAAGAGGAACTGT 342
QY 542 AAAATGCCAGAGAGGGAATCAAGGTCTTAAAGGAAGATGAAGCTGAAACAAAC 601
Db 343 AAAATGCCAGAGAGGGAATCAAGGTCTTAAAGGAAGATGAAGCTGAAACAAAC 402

QY 602 ACAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAATAATTTT 655
Db 403 GCAAGCTGGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAATAATTTT 456

RESULT 15
US-09-702-705-792
; Sequence 792, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 792
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-792

Query Match 26.6%; Score 178; DB 4; Length 461;
Best Local Similarity 73.9%; Pred. No. 3,9e-47;
Matches 260; Conservative 0; Mismatches 75; Indels 17; Gaps 2;

QY 319 GTCCAGGACCCACTGATGAAGAGCCTTAAGAGAGAGAAACCCACTTAAAGTCGAA 378
Db 101 GTCCAGGAGCCAGTAAATGGAGAGCCCAAAAGAGAAACAGCAGCTGAAGTCGGA 160
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGTGCGAGTGCAGATTC----- 428
Db 161 TCCTACACCTGGGAGAGAGATGATCAGGTGCGAGTGCAGATTC----- 220
QY 429 -----AAGTGCCTGACCTGGAAGCGGATCTCCAGGAGCTATGTTCAGACAAAGACTGGG 482
Db 221 GAAGGGAATGCGGACATGGAAGGTGATCTGCAAGAGCTGCATCAGTCAAAACACCGGG 280
QY 483 ATGATGTGAAGGT-GGTACTGATGTCAAGGGGAAGTCTTACCAGAGAGAGACTTT 541
Db 281 ATAAATCTGGATTTGGGTTCCGGCTCAAGGTGAAGATAATACCTTAAAGAGGAACTGT 340
QY 542 AAAATGCCAGAGAGGGAATCAAGGTCTTAAAGGAAGATGAAGCTGAAACAAAC 601
Db 341 AAAATGCCAGAGAGGGAATCAAGGTCTTAAAGGAAGATGAAGCTGAAACAAAC 400
QY 602 ACAACTGTTTTATATTAGATATTTTACTTTAAAGAGTCTTAATAATTTT 653
Db 401 GCAAGCTGGTTTTATATTAGATATTTTACTTTAAAGTCTTAATAATTTT 452

RESULT 16
US-09-736-457-792
; Sequence 792, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick

STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/531,662B
FILING DATE: 21-September-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/370,648
FILING DATE: 10-January-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250,162
FILING DATE: 27-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,039
FILING DATE: 22-July-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5858689man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5323.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-531-662B-17

Query Match 25.4%; Score 169.8; DB 2; Length 532;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;

Qy 217 GTGAATATGATGGCGAGAGATCAAC--ATATAGCCTAGGCCAGGAGAGTTT 273
Db 68 GTGAATATGATGGCGAGAGATCAAC--ATATAGCCTAGGCCAGGAGAGTTT 127

Qy 274 ACAGCCTCTGAGCTGATGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCAC 333
Db 128 ACAGCCTCTGAGCTGATGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCAC 184

Qy 334 TGATGAGAGCCTAAAGAGAGAACCCACCACTAAAGTCGGAATCTTACCTGATCA 393
Db 185 ACCAGCAACCTGAAAGAGAGAACCCACCACTAAAGTCGGAATCTTACCTGATCA 244

Qy 394 GAAGAGAGAGATGATCAGGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCT 453
Db 245 GGAGGAGAGATG--AGGAGCATCTGAGTCAAGGSCCGAGCTGAGCTGATAG 301

Qy 454 CCAGGAGCTATGTCAGAACAGACTGGGGATGATGTAAGTGGTACTGATGTCAGG 513
Db 302 CCAGGAGAGAGGTCACCCAGACTGGGTGTGATGTAAGTGGTCTGATGGCAGGA 361

Qy 514 GAAGATCTACCAAAAGCAGAGCACTTTAAATGCCAGAGAGGTCGAGGGAATCA 573
Db 362 GATGACCCGCAATCCAGAGAGGTCGAGGTCGAGAGAGGTCGAGAGGTCGAG 421

Qy 574 GGTGTAAGGAGATAGCTGAAACCAACAACTGTTTTTATATATATATATTTTACT 633
Db 422 GTGTTAAAGAGGAGGAGGCTGAAATGATGAGGCTGCTCTATGTTGGAATTTG 481

Qy 634 AAGAGTCT 642
Db 482 TAAATTTCT 490

RESULT 19
US-08-669-161A-17
Sequence 17, Application US/08669161A
Patent No. 6013481
GENERAL INFORMATION:
APPLICANT: Debacker, Olivier; Van den Eynde,
APPLICANT: Benoit; Boon-Falleur, Thierry
TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pelfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/669,161A
FILING DATE: 24-June-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/531,662
FILING DATE: 21-September-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/370,648
FILING DATE: 10-January-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250,162
FILING DATE: 27-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,039
FILING DATE: 22-July-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6013481man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5443
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-669-161A-17

Query Match 25.4%; Score 169.8; DB 3; Length 532;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;

Qy 217 GTGAATATGATGGCGAGAGATCAAC--ATATAGCCTAGGCCAGGAGAGTTT 273
Db 68 GTGAATATGATGGCGAGAGATCAAC--ATATAGCCTAGGCCAGGAGAGTTT 127

Qy 274 ACAGCCTCTGAGCTGATGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCAC 333
Db 128 ACAGCCTCTGAGCTGATGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCAC 184

Qy 334 TGATGAGAGCCTAAAGAGAGAACCCACCACTAAAGTCGGAATCTTACCTGATCA 393
Db 185 ACCAGCAACCTGAAAGAGAGAACCCACCACTAAAGTCGGAATCTTACCTGATCA 244

Qy 394 GAAGAGAGAGATGATCAGGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCT 453
Db 245 GGAGGAGAGATG--AGGAGCATCTGAGTCAAGGSCCGAGCTGAGCTGATAG 301

Qy 454 CCAGGAGCTATGTCAGAACAGACTGGGGATGATGTAAGTGGTACTGATGTCAGG 513
Db 302 CCAGGAGAGAGGTCACCCAGACTGGGTGTGATGTAAGTGGTCTGATGGCAGGA 361

Qy 514 GAAGATCTACCAAAAGCAGAGCACTTTAAATGCCAGAGAGGTCGAGGGAATCA 573
Db 362 GATGACCCGCAATCCAGAGAGGTCGAGGTCGAGAGAGGTCGAGAGGTCGAG 421

Qy 574 GGTGTAAGGAGATAGCTGAAACCAACAACTGTTTTTATATATATATATTTTACT 633
Db 422 GTGTTAAAGAGGAGGAGGCTGAAATGATGAGGCTGCTCTATGTTGGAATTTG 481

Qy 634 AAGAGTCT 642
Db 482 TAAATTTCT 490

Db 245 GGAGGAGAGGATG---AGGAGCATCTGCAGTCAAGGCCGAGACCTGAGCTGATAG 301
QY 454 CCAGGAGCTATGTACAGACAAAGCTGGGATGGATGTGAAGTGTGCTACTGATGTCAAGG 513
Db 302 CCAGGAAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGTGTCTCTGATGGGCAGGA 361
QY 514 GAAGATTCTACCAAGACAGACACTTTAAATGCCAGAGCAGGTGAAGGAGAAATCACA 573
Db 362 GATGGACCCGCCAAATCCAGAGAGGTGAAGCCCTGAAGAGGTGAAGCAATCACA 421
QY 574 GGTTTAAAGGAGATAGCTGAACACAACTGTTTTATATTAGATATTTTACTTT 633
Db 422 GTGTTAAAGAGGACGTTGAATGATGAGGCTGCTCTATGTTGGAATTTGTTAT 481
QY 634 AAGAGCTCT 642
Db 482 TAAATTTCT 490

RESULT 20
US-08-602-039-17
; Sequence 17, Application US/08602039
; Patent No. 6063001
; GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry;
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,039
; FILING DATE: 15-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-JANUARY-1995
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-JULY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6069001man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5323.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-602-039-17

Query Match 25.4%; Score 169.8; DB 3; Length 532;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;

QY 217 GTGAATATCAGTTGGCGAGAGATCAAC---ATATAGGCTAGGCCAAGAGAGAGTTT 273
Db 68 GTGAATATCAGTTGGCGAGAGATCGACTTATTATTGGCCCTAGACCAAGGCGCTAATGT 127
QY 274 ACAGCTCTCTGAGCTGATTGGGGCTATGCTTACTTGGCTCCCTTTTGTCCAGGAACCCAC 333
Db 128 ACAGCTCTGAAAGTGAATTGGGCTATGCGGCC---CGAGCAGTTCAAGTCATGAAGTGA 184
QY 334 TGATGAAGAGCTTAAAGAGAGAAACCCACCTAAAGTCCGAATCCCTACACCTGATCA 393
Db 185 ACCAGCAACACCTGAAGAGAGGAAACCCAGCAACTCAACGTCAGGATCCTCGAGCTGCTCA 244
QY 394 GAAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGATTCAGCTGCCTGACCTGAAAGCCGATCT 453
Db 245 GGAGGGAGAGGATG---AGGAGCATCTGCAGGTCAAGGGCCGAAAGCTCAAGCTGATAG 301
QY 454 CCAGGAGCTATGTACAGACAAAGACTGGGATGGATGTGAAGTGTGCTACTGATGTCAAGG 513
Db 302 CCAGGAAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGTGTCTCTATGTTGGAATTTGTTAT 481
QY 514 GAAGATTCTACCAAGACAGACACTTTAAATGCCAGAGCAGGTGAAGGAGAAATCACA 573
Db 362 GATGGACCCGCCAAATCCAGAGAGGTGAAGCCCTGAAGAGGTGAAGCAATCACA 421
QY 574 GGTTTAAAGGAGATAGCTGAACACAACTGTTTTATATTAGATATTTTACTTT 633
Db 422 GTGTTAAAGAGGACGTTGAATGATGAGGCTGCTCTATGTTGGAATTTGTTAT 481
QY 634 AAGAGCTCT 642
Db 482 TAAATTTCT 490

RESULT 21
US-08-370-648-18
; Sequence 18, Application US/08370648
; Patent No. 5648226
; GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry;
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,648
; FILING DATE: 10-JANUARY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-JULY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5648226man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5323.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200

```
TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 539 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-370-648-18

Query Match      25.4%; Score 169.8; DB 1; Length 539;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;

QY 217 GTCAATATGATGGCGAGAGATCAAC---ATATAGCCTAGGCCAAGAAGATT 273
Db 75 GTCAATATGATGGCGAGAGATCGACTATTATGGCTAGACCAAGCGCTATGT 134
QY 274 ACAGCCTCTGAGCTGATGGGCTATGCTTACTGGCTCCCTTTGCCAGGAACCCAC 333
Db 135 ACAGCCTCTGAGCTGATGGGCTATGCTTACTGGCTCCCTTTGCCAGGAACCCAC 191
QY 334 TGATGAGAGCCTAAGAGAGAAACCCACCCTAAAGTCGGAATCCCTACACCTGATCA 393
Db 192 ACAGCAACACCTGAGAGAGGGAACCAAGCACTCAAGTCAAGATCTCGAGCTGCTCA 251
QY 394 GAAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCT 453
Db 252 GGAGGGAGAGATG---AGGAGCATCTGCAGGTCAAGGGCCGAAGCCTGAAGCTGATAG 308
QY 454 CCAGGAGTATGTCAGACAAAGACTGGGGATGGATGTGAAGTGGTACTGATCTCAAGGG 513
Db 309 CCAGGAAACAGGGTCAACCAAGCTGGGGTGTGAGTGTGAAGATGGTCTGATGGCAGGA 368
QY 514 GAAGATTCTACCAAAAGCAGAGCACTTTAAAGTCCAGAGAGAGGTGAAGGAAATCAACA 573
Db 369 GGTGACCCGCCAATCCAGAGGAGGTGAAAGCCCTGAAGAGGTGAAAGCAATCAACA 428
QY 574 GGTTTAAGGAGATAGCTGAACACACAACTGTTTATATAGATTTTACTTT 633
Db 429 GTGTTAAAGAGAGACAGCTTGAATGATGCAAGCTGCTCTATGTTGGAATTTGTTCA 488
QY 634 AAAGAGTCT 642
Db 489 TAAATTTCT 497

RESULT 22
US-08-531-662B-18
; Sequence 18, Application US/08531662B
; Patent No. 5858689
; GENERAL INFORMATION:
; APPLICANT: Van der Bruggen, Pierre;
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated, Truncated Nucleic
; TITLE OF INVENTION: Acid Molecules Which Code For GAGE Tumor
; TITLE OF INVENTION: Rejection Antigen, The Tumor Rejection Antigen,
; TITLE OF INVENTION: and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felle & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,662B
```

```
FILING DATE: 21-September-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-January-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-May-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-July-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5858689man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5323.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 539 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-531-662B-18

Query Match      25.4%; Score 169.8; DB 2; Length 539;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;

QY 217 GTCAATATGATGGCGAGAGATCAAC---ATATAGCCTAGGCCAAGAAGATT 273
Db 75 GTCAATATGATGGCGAGAGATCGACTATTATGGCTAGACCAAGCGCTATGT 134
QY 274 ACAGCCTCTGAGCTGATGGGCTATGCTTACTGGCTCCCTTTGCCAGGAACCCAC 333
Db 135 ACAGCCTCTGAGCTGATGGGCTATGCTTACTGGCTCCCTTTGCCAGGAACCCAC 191
QY 334 TGATGAGAGCCTAAGAGAGAAACCCACCCTAAAGTCGGAATCCCTACACCTGATCA 393
Db 192 ACAGCAACACCTGAGAGAGGGAACCAAGCACTCAAGTCAAGATCTCGAGCTGCTCA 251
QY 394 GAAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCT 453
Db 252 GGAGGGAGAGATG---AGGAGCATCTGCAGGTCAAGGGCCGAAGCCTGAAGCTGATAG 308
QY 454 CCAGGAGTATGTCAGACAAAGACTGGGGATGGATGTGAAGTGGTACTGATCTCAAGGG 513
Db 309 CCAGGAAACAGGGTCAACCAAGCTGGGGTGTGAGTGTGAAGATGGTCTGATGGCAGGA 368
QY 514 GAAGATTCTACCAAAAGCAGAGCACTTTAAAGTCCAGAGAGAGGTGAAGGAAATCAACA 573
Db 369 GGTGACCCGCCAATCCAGAGGAGGTGAAAGCCCTGAAGAGGTGAAAGCAATCAACA 428
QY 574 GGTTTAAGGAGATAGCTGAACACACAACTGTTTATATAGATTTTACTTT 633
Db 429 GTGTTAAAGAGAGACAGCTTGAATGATGCAAGCTGCTCTATGTTGGAATTTGTTCA 488
QY 634 AAAGAGTCT 642
Db 489 TAAATTTCT 497

RESULT 23
US-08-669-161A-18
; Sequence 18, Application US/08669161A
; Patent No. 6013481
; GENERAL INFORMATION:
; APPLICANT: Debacter, Olivier; Van den Eynde,
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
```

NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felle & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/669,161A
FILING DATE: 24-June-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/531,662
FILING DATE: 21-September-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/370,648
FILING DATE: 10-January-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250,162
FILING DATE: 27-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,039
FILING DATE: 22-July-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6013481man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5443
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-669-161A-18
Query Match 25.4%; Score 169.8; DB 3; Length 539;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;
QY 217 GTGAAATATGAGTTGGCGAGGAGATCAAC---ATATAGGCTTAGGCCAAGAGAGTTT 273
DB 75 GTGAAATATGAGTTGGCGAGGAGATCAAC---ATATAGGCTTAGGCCAAGAGAGTTT 134
QY 274 ACAGCCTCTGAGCTGATTGGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCAC 333
DB 135 ACAGCCTCTGAGCTGATTGGGGCTATGCGGCC---CGAGCAGTTTCAGTGATGAAGTGA 191
QY 394 GAAGAGAGATGATCAGGGTGCAGCTGAGATCAAGTGCCTCACTGGAAGCCGATCT 453
DB 252 GGAGGGAGAGATG---AGGAGCATCTGCAAGGCTCAAGGGCCGAGCTGAAGCTGATAG 308
QY 454 CCAGAGCTATGTACAGACAAAGATGCGGATGGAATGGAAGTGGTACTGATGTCAAGGG 513
DB 309 CCAGAAACAGGTCACCCACAGACTGGGTGTGAGTGTGAAGATGTTCTGATGGCAGGA 368
QY 514 GAAGATTCTCAAAAGCAGAGCACTTTAAATGTCAGAGCAGGTGAAGGAAATCACA 573
DB 369 GGTGACCCGCCAAATCCAGAGGAGGTGAAACCGCTGAGAGAGGTGAAAAGCAATCACA 428
QY 574 GGTTTAAGGAGATGAAGCTGAACACACAACTGTTTTTATATTAGATATTTTACTTT 633
US-08-602-039-18
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,039
FILING DATE: 15-FEB-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/370,648
FILING DATE: 10-JANUARY-1995
APPLICATION NUMBER: 08/250,162
FILING DATE: 27-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,039
FILING DATE: 22-JULY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6069001man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5323.2
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-602-039-18
Query Match 25.4%; Score 169.8; DB 3; Length 539;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;
QY 217 GTGAAATATGAGTTGGCGAGGAGATCAAC---ATATAGGCTTAGGCCAAGAGAGTTT 273
DB 75 GTGAAATATGAGTTGGCGAGGAGATCAAC---ATATAGGCTTAGGCCAAGAGAGTTT 134
QY 274 ACAGCCTCTGAGCTGATTGGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCAC 333
DB 135 ACAGCCTCTGAGCTGATTGGGGCTATGCGGCC---CGAGCAGTTTCAGTGATGAAGTGA 191
QY 334 TGATGAAGAGCCCTAAAGAGAGAGAAACCCACTCACTAAAGTCCGATCTCAGCTGATCA 393
DB 192 ACAGCAACACCTGAGAGAGAGAGAAACCCAGCAACTCAACGTGAGTTTCAGTGATGAAGTGA 251
US-08-602-039-18
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,039
FILING DATE: 15-FEB-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/370,648
FILING DATE: 10-JANUARY-1995
APPLICATION NUMBER: 08/250,162
FILING DATE: 27-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,039
FILING DATE: 22-JULY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6069001man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5323.2
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-602-039-18
Query Match 25.4%; Score 169.8; DB 3; Length 539;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;
QY 217 GTGAAATATGAGTTGGCGAGGAGATCAAC---ATATAGGCTTAGGCCAAGAGAGTTT 273
DB 75 GTGAAATATGAGTTGGCGAGGAGATCAAC---ATATAGGCTTAGGCCAAGAGAGTTT 134
QY 274 ACAGCCTCTGAGCTGATTGGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCAC 333
DB 135 ACAGCCTCTGAGCTGATTGGGGCTATGCGGCC---CGAGCAGTTTCAGTGATGAAGTGA 191
QY 334 TGATGAAGAGCCCTAAAGAGAGAGAAACCCACTCACTAAAGTCCGATCTCAGCTGATCA 393
DB 192 ACAGCAACACCTGAGAGAGAGAGAAACCCAGCAACTCAACGTGAGTTTCAGTGATGAAGTGA 251

QY 394 GAAGAGAGAGATGATCAGGTCAGCTGAGATTCAGTCTGACCTGGAAGCCGATCT 453
DB 252 GGAGGGAGAGGATG---AGGAGCATCTGCAGGTCAAGGCCGGAAGCCCTGGAAGCTGTAG 308
QY 454 CCAGGAGCTATGTCAGACAAAGACTGGGATGGATGTAAGGTGGTACTGATGTCAAGGG 513
DB 309 CCAGGAAACAGGTCACCCACAGACTGGGTGTGATGTGAAGTGTCTCTGATGGCGCAGA 368
QY 514 GAAGATTCACCAAAAGCAGAGACACTTTAAATGCCAGAGCAGGTCGAAGGGAATCACA 573
DB 369 GGTGAGACCCGCCCAATCCAGAGAGGTGAAACGCCCTGAAGAAGGTGAAAGCAATCACA 428
QY 574 GGTTTAAAGCAAGATAGCTGAAACACACACTGTTTTATATTAGATATTTTACTTT 633
DB 429 GTGTTAAAGAACACAGCTTGAATGATGACGCTCTCTATGTTGGAATTTGTTCA 488
QY 634 AAAGAGTCT 642
DB 489 TAAATTTCT 497

RESULT 25

US-08-370-648-16
; Sequence 16, Application US/08370648
; Patent No. 5648226
; GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry;
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,648
; FILING DATE: 10-JANUARY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER:
; FILING DATE: 27-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-JULY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5648226man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5323.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 540 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-370-648-16
Query Match 25.4%; Score 169.8; DB 1; Length 540;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;

QY 217 GTGAATATTCAGTTGGCGGAGAGATCAAC---ATATAGCCTAGGCCAGGAAAGATTT 273
DB 76 GTGAATATTCAGTTGGCGGAGAGATCGACCTATTATTGGCCCTAGACCAAGCGCTATGT 135
QY 274 ACAGCTCTCTGAGCTGATTTGGGCTATGCTTACTGGCTCCCTTTGTCOCAGAAACCCAC 333
DB 136 ACAGCTCTCTGAAATGATTGGGCTATGCGGCC---CGAGCAGTTCAAGTATGAAGTGA 192
QY 334 TGATGAAGAGCCTTAAAGAGAGAAACACCCACTTAAAGTCGGAATCCCTACACCTGATCA 393
DB 193 ACCAGCAACACCTTAAAGAGAGAAACACCACTCAACAGTCAGAGATCCTCAGCTGCTCA 252
QY 394 GAAAGAGAGAGATGATCAGGTCGAGCTGAGATTCAGTTCCTGACCTGGAAGCCGATCT 453
DB 253 GGAGGAGAGGATG---AGGAGCATCTGCAGGTCAAGGCCGGAAGCCTGAAGCTGATAG 309
QY 454 CCAGAGCTATGTCAGACAAAGACTGGGATGGATGTGAAGGTGGTACTGATGTCAAGGG 513
DB 310 CCAGGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGGTCTCTATGTTGGA 369
QY 514 GAAGATTCACCAAAAGCAGAGCACTTTAAATGCCAGAGCAGGTCGAAGGGAATCACA 573
DB 370 GATGACCCGCCCAATCCAGAGAGGTGAAACGCCCTGAAGAGGTGAAAGCAATCACA 429
QY 574 GGTTTAAAGCAAGATAGCTGAAACACACACTGTTTTATATTAGATATTTTACTTT 633
DB 430 GTGTTAAAGAAAGGCACGCTTGAATGATGACGCTCTCTATGTTGGAATTTGTTCA 489
QY 634 AAAGAGTCT 642
DB 490 TAAATTTCT 498

RESULT 26

US-08-531-662B-16
; Sequence 16, Application US/08531662B
; Patent No. 5858689
; GENERAL INFORMATION:
; APPLICANT: Van der Bruggen, Pierre;
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated, Truncated Nucleic
; TITLE OF INVENTION: Acid Molecules Which Code For GAGE Tumor
; TITLE OF INVENTION: Rejection Antigen, The Tumor Rejection Antigen,
; TITLE OF INVENTION: and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,662B
; FILING DATE: 21-September-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-January-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-May-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-July-1993
; ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 585868man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5323.3
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-531-662B-16

Query Match 25.4%; Score 169.8; DB 2; Length 540;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;

QY 217 GTGAATATGAGTGGCGAGAGATCAAC---ATATAGCCTAGCGCCAAAGAAAGTTT 273
DB GTGAATATGAGTGGCGAGAGATCAACCTATTATTGGCTTAGACCAAGGCGCTATGT 135
QY 274 ACAGCCTCTGAGTGGCGAGAGATCAAC---ATATAGCCTAGCGCCAAAGAAAGTTT 333
DB ACAGCCTCTGAGTGGCGAGAGATCAACCTATTATTGGCTTAGACCAAGGCGCTATGT 135
QY 334 TGATGAAGAGCCTAAAGAGAGAGAAACCAACCTAAAGTGGAAATCTTACACCTGATCA 393
DB TGATGAAGAGCCTAAAGAGAGAGAAACCAACCTAAAGTGGAAATCTTACACCTGATCA 393
QY 193 ACCAGCAACACTGAAGAGAGAGAAACCAACCTAAAGTGGAAATCTTACACCTGATCA 252
DB ACCAGCAACACTGAAGAGAGAGAAACCAACCTAAAGTGGAAATCTTACACCTGATCA 252
QY 394 GAAGAGAGAGATGATGAGTGGCGAGAGATCAAC---ATATAGCCTAGCGCCAAAGAAAGTTT 453
DB GAAGAGAGAGATGATGAGTGGCGAGAGATCAACCTATTATTGGCTTAGACCAAGGCGCTATGT 135
QY 454 CCAGGAGCTATGTCAGACAAAGAGAGAGAAACCAACCTAAAGTGGAAATCTTACACCTGATCA 393
DB CCAGGAGCTATGTCAGACAAAGAGAGAGAAACCAACCTAAAGTGGAAATCTTACACCTGATCA 393
QY 370 GATGAGAGAGATGATGAGTGGCGAGAGATCAAC---ATATAGCCTAGCGCCAAAGAAAGTTT 429
DB GATGAGAGAGATGATGAGTGGCGAGAGATCAACCTATTATTGGCTTAGACCAAGGCGCTATGT 135
QY 574 GGTAAAGAGAGATGATGAGTGGCGAGAGATCAAC---ATATAGCCTAGCGCCAAAGAAAGTTT 633
DB GGTAAAGAGAGATGATGAGTGGCGAGAGATCAACCTATTATTGGCTTAGACCAAGGCGCTATGT 135
QY 634 AAAGAGTCT 642
DB TAAATTTCT 498

RESULT 27
US-08-669-161A-16
Sequence 16, Application US/08669161A
Patent No. 6013481
GENERAL INFORMATION:
APPLICANT: DeBacker, Olivier; Van den Eynde,
APPLICANT: Benoit; Boon-Talieu, Thierry
TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/669,161A
FILING DATE: 24-June-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/531,662
FILING DATE: 21-September-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/370,648
FILING DATE: 10-January-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250,162
FILING DATE: 27-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,039
FILING DATE: 22-July-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6013481man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5443
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-669-161A-16

Query Match 25.4%; Score 169.8; DB 3; Length 540;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;

QY 217 GTGAATATGAGTGGCGAGAGATCAAC---ATATAGCCTAGCGCCAAAGAAAGTTT 273
DB GTGAATATGAGTGGCGAGAGATCAACCTATTATTGGCTTAGACCAAGGCGCTATGT 135
QY 274 ACAGCCTCTGAGTGGCGAGAGATCAAC---ATATAGCCTAGCGCCAAAGAAAGTTT 333
DB ACAGCCTCTGAGTGGCGAGAGATCAACCTATTATTGGCTTAGACCAAGGCGCTATGT 135
QY 334 TGATGAAGAGCCTAAAGAGAGAGAAACCAACCTAAAGTGGAAATCTTACACCTGATCA 393
DB TGATGAAGAGCCTAAAGAGAGAGAAACCAACCTAAAGTGGAAATCTTACACCTGATCA 393
QY 193 ACCAGCAACACTGAAGAGAGAGAAACCAACCTAAAGTGGAAATCTTACACCTGATCA 252
DB ACCAGCAACACTGAAGAGAGAGAAACCAACCTAAAGTGGAAATCTTACACCTGATCA 252
QY 394 GAAGAGAGAGATGATGAGTGGCGAGAGATCAAC---ATATAGCCTAGCGCCAAAGAAAGTTT 453
DB GAAGAGAGAGATGATGAGTGGCGAGAGATCAACCTATTATTGGCTTAGACCAAGGCGCTATGT 135
QY 454 CCAGGAGCTATGTCAGACAAAGAGAGAGAAACCAACCTAAAGTGGAAATCTTACACCTGATCA 393
DB CCAGGAGCTATGTCAGACAAAGAGAGAGAAACCAACCTAAAGTGGAAATCTTACACCTGATCA 393
QY 310 CCAGGAGAGAGATGATGAGTGGCGAGAGATCAAC---ATATAGCCTAGCGCCAAAGAAAGTTT 429
DB CCAGGAGAGAGATGATGAGTGGCGAGAGATCAACCTATTATTGGCTTAGACCAAGGCGCTATGT 135
QY 514 GAAGATTTTACCAAGAGAGAGATCAAC---ATATAGCCTAGCGCCAAAGAAAGTTT 573
DB GAAGATTTTACCAAGAGAGAGATCAACCTATTATTGGCTTAGACCAAGGCGCTATGT 135
QY 370 GATGAGAGAGATGATGAGTGGCGAGAGATCAAC---ATATAGCCTAGCGCCAAAGAAAGTTT 429
DB GATGAGAGAGATGATGAGTGGCGAGAGATCAACCTATTATTGGCTTAGACCAAGGCGCTATGT 135
QY 574 GGTAAAGAGAGATGATGAGTGGCGAGAGATCAAC---ATATAGCCTAGCGCCAAAGAAAGTTT 633
DB GGTAAAGAGAGATGATGAGTGGCGAGAGATCAACCTATTATTGGCTTAGACCAAGGCGCTATGT 135
QY 430 GTGTTAAAGAGAGAGATGATGAGTGGCGAGAGATCAAC---ATATAGCCTAGCGCCAAAGAAAGTTT 489
DB GTGTTAAAGAGAGAGATGATGAGTGGCGAGAGATCAACCTATTATTGGCTTAGACCAAGGCGCTATGT 135
QY 634 AAAGAGTCT 642
DB TAAATTTCT 498

RESULT 28
US-08-602-039-16
Sequence 16, Application US/08602039
Patent No. 6069001

GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry;
; TITLE OF INVENTION: Isolated, truncated Nucleic Acid
; MOLECULES, Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,039
; FILING DATE: 15-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-JULY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No 6069001man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5323.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 540 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-602-039-16

Query Match 25.4%; Score 169.8; DB 3; Length 540;
Best Local Similarity 67.1%; Pred. No. 1.8e-44;
Matches 288; Conservative 0; Mismatches 132; Indels 9; Gaps 3;
QY 217 GTGAAATATGAGTTGGCGAGGAGATCAAC---ATATAGGCTTAGGCCAAGAGAGTTT 273
DB 76 GTGAAATATGAGTTGGCGAGGAGATCGACCTATTATTGGCTTAGACCAAGCGCTATGT 135
QY 274 ACAGCTCTGAGCTGATTTGGGCTATGTTACTGGCTCCCTTTGTCCAGGAAACCCAC 333
DB 136 ACAGCTCTGAAATGATTTGGGCTATGGGCC---CGAGCAGTTGAGTGAAGTGA 192
QY 334 TGATGAAGACCTTAAAGAGAGAGAAACCCACCTAAAGTTCGGAATCTTACACCTGATCA 393
DB 193 ACCAGCAACACCTGAGAGAGGGGAAACAGCACTCAACGTCAGGATCTTGCAGCTGCTCA 252
QY 394 GAAGAGAGAGATGATCAGGGTCAGCTGAGATTCAGTTCAGTTCAGTTCAGTTCAGTTC 453
DB 253 GGAGGGAGAGATG---AGGGAGATCTGAGGTCAGGGCCGAGGCTGAGCTGATAG 309
QY 454 CCAGGAGCTATGTCAGCAAAAGACTGGGGATGGATGTGAAGTGTGAGTGTGAGTGTGAG 513
DB 310 CCAGGACAGGGTACCCACAGACTGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 369
QY 514 GAAGATTTACCAAAACAGAGCACTTTAAATGCCAGAGAGAGGTCGAGGGAATCACA 573
DB 370 GATGGACCCGCAATCCAGAGGAGGTGAAACCGCTGAAAGAGGTTGAAAGCAATCACA 429

QY 574 GGTTTAAAGGAGATGAGCTCAAAACACAAACAACTGTTTTATATATAGATATTTTACTTT 633
DB 430 GTGTTAAAGAGGACGCTTGAATGATGACGGTCTCTTATGTTGAAATTTTTCAT 489
QY 634 AAGAGTCT 642
DB 490 TAAATTTCT 498
RESULT 29
US-09-163-748C-2
; Sequence 2, Application US/09163748C
; Patent No. 6509172
; GENERAL INFORMATION:
; APPLICANT: DeBacker, Olivier
; APPLICANT: Van den Eynde, Benoit
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Encode Member Of The Gage
; TITLE OF INVENTION: Family, The Proteins Encoded, And Uses Thereof
; FILE REFERENCE: LUD 5558
; CURRENT APPLICATION NUMBER: US/09/163,748C
; CURRENT FILING DATE: 1998-09-30
; NUMBER OF SEQ ID NOS: 27
; SEQ ID NO 2
; LENGTH: 526
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; US-09-163-748C-2

Query Match 25.1%; Score 168.2; DB 4; Length 526;
Best Local Similarity 66.9%; Pred. No. 5.7e-44;
Matches 287; Conservative 0; Mismatches 133; Indels 9; Gaps 3;
QY 217 GTCAATATGAGTTGGCGAGGAGATCAAC---ATATAGGCTTAGGCCAAGAGAGTTT 273
DB 48 GTCAATATGAGTTGGCGAGGAGATCGACCTATTATTGGCTTAGACCAAGCGCTATGT 107
QY 274 ACAGCTCTGAGCTGATTTGGGCTATGTTACTGGCTCCCTTTGTCCAGGAAACCCAC 333
DB 108 ACAGCTCTGAAATGATTTGGGCTATGGGCC---CGAGCAGTTTCAGTGAAGTGA 164
QY 334 TGATGAAGACCTTAAAGAGAGAGAAACCCACCTAAAGTTCGGAATCTTACACCTGATCA 393
DB 165 ACCAGCAACACCTGAGAGAGGGGAAACCAAGCACTCAAGTCAGGATCTTGCAGCTGCTCA 224
QY 394 GAAGAGAGAGATGATCAGGGTCAGCTGAGATTCAGTTCAGTTCAGTTCAGTTCAGTTC 453
DB 225 GGAGGGAGAGATG---AGGGAGATCTGAGGTCAGGGCCGAGGCTGAAAGCTCATAG 281
QY 454 CCAGGAGCTATGTCAGCAAAAGACTGGGATGGATGTGAAGTGTGATGTCAAGGG 513
DB 282 CCAGGAAACAGGGTCCACCAGACTGGTGTGAGTGTGAAGATGTTCTGATGGCAGGA 341
QY 514 GAAGATTTACCAAAACAGAGCACTTTAAATGCCAGAGAGGTCGAGGTAAGGGAATCACA 573
DB 342 GATGACCCGCCAATCCAGAGGAGGTGAAACCGCTGAAAGAGGTGAAAGCAATCACA 401
QY 574 GGTTTAAAGGAGATGAGCTGAAACACACAACTGTTTTATATAGATATTTTACTTT 633
DB 402 GTGTTAAAGAGGACGCTTGAATGATGACGGCTGCTCTCTATGTTGAAATTTTGTTCAT 461
QY 634 AAGAGTCT 642
DB 462 TAAATTTCT 470

RESULT 30
US-08-370-648-15
; Sequence 15, Application US/08370648
; Patent No. 5648226
; GENERAL INFORMATION:

APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
APPLICANT: Boon-falleur, Thierry;
TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370,648
FILING DATE: 10-JANUARY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250,162
FILING DATE: 27-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,039
FILING DATE: 22-JULY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5648226man D.
REGISTRATION NUMBER: LUD 5323.2
REFERENCE/DOCKET NUMBER: LUD 5323.2
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 560 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-370-648-15
Query Match 24.2%; Score 162.2; DB 1; Length 560;
Best Local Similarity 62.4%; Pred. No. 4.9e-42; Indels 6; Gaps 2;
Matches 289; Conservative 0; Mismatches 168; Indels 6; Gaps 2;
Qy 180 ATAGGTCCGCGAGCCAGTCTCCAGGAAGTGAATATGTAATATGAGTTGGCGAGAA 239
Db 62 ATAGGTGATCTCTCTCGCAACTCATATTCACACAGATGAATCTCAGTAGAGGAAAT 121
Qy 240 GATCAACATATAGGCTTAGCCCAAGAAAGTTTACAGCTCTCCTGAGTTGGGCTA 299
Db 122 CGACCTATTATTGGCTTAGCCCAAGCGCTATGTACAGCTCTCCTGAGTTGGGCTA 181
Qy 300 TGCTTACTGGCTCCCTTTGTGCCAGGAACCCACTGATGAGAGCTTAAAGAGAGAAC 359
Db 182 TGGCGCCGAG--CAGTTCAGTGAAGTGAAGAACCCAGCAACACTGAAGAGGGAAC 238
Qy 360 CACCACATAAGTGGATCTACCTGATCAGAGAGAGATGATCAGGTTGCG 419
Db 239 CAGCAACTCAACGTCAGATCTCAGCTGCTCAGGAGGAGAGATG--AGGAGCAT 295
Qy 420 CTGAGATTCAAGTGCCTCAGCTGGAAGCCGATCTCCAGGAGCTATGTGACAGCAAACTG 479
Db 296 CTGAGGTCAGGGCCGAGGCTGAGCTGATAGCCAGGACAGGCTCACCACAGACTG 355
Qy 480 GGGATGATGAGGTTGGTATCTGATGTCAGGGAAGATTTCTACAAAGAGAGACT 539
Db 356 GGTGTGATGTGAAGATGGTCTGATGGGAGGAGATGGACCCGCCAAATCCAGAGAGG 415
Qy 540 TTAATAATCCAGAGCAGTGAAGGGAATCAGAGTTTAAAGGAAGATGATGTAACA 599

Db 416 TGAACGCGCTGAGAGAGGTGAAGCAATCAGAGTGTAAAGAGGACGCTTGAATG 475
Qy 600 ACACAACTGTTTTATATATAGATATTTTACTTTAAAGAGTCT 642
Db 476 ATGAGGCTGCTCTCTATGTTGGAATTTGTTCAATAAAATCT 518
RESULT 31
US-08-531-662B-15
Sequence 15, Application US/08531662B
Patent No. 5858689
GENERAL INFORMATION:
APPLICANT: Van der Bruggen, Pierre;
APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
APPLICANT: Boon-falleur, Thierry
TITLE OF INVENTION: Isolated, Truncated Nucleic
TITLE OF INVENTION: Acid Molecules Which Code For GAGE Tumor
TITLE OF INVENTION: Rejection Antigen, The Tumor Rejection Antigen,
TITLE OF INVENTION: and Uses Thereof
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/531,662B
FILING DATE: 21-September-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/370,648
FILING DATE: 10-January-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250,162
FILING DATE: 27-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,039
FILING DATE: 22-July-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5858689man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5323.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 560 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-531-662B-15

Query Match 24.2%; Score 162.2; DB 2; Length 560;
Best Local Similarity 62.4%; Pred. No. 4.9e-42; Indels 6; Gaps 2;
Matches 289; Conservative 0; Mismatches 168; Indels 6; Gaps 2;
Qy 180 ATAGGTCCGCGAGCCAGTCTCCAGGAAGTGAATATGTAATATGAGTTGGCGAGAA 239
Db 62 ATAGGTGATCTCTCTCGCAACTCATATTCACACAGATGAATCTCAGTAGAGGAAAT 121
Qy 240 GATCAACATATAGGCTTAGCCCAAGAAAGTTTACAGCTCTCCTGAGTTGGGCTA 299
Db 122 CGACCTATTATTGGCTTAGCCCAAGCGCTATGTACAGCTCTCCTGAGTTGGGCTA 181
Qy 300 TGCTTACTGGCTCCCTTTGTGCCAGGAACCCACTGATGAGAGCTTAAAGAGAGAAC 359

182 TCGGCCCCGAG--CAGTTTCAGTGATGAAGTGGACACGACACCTTGAGAGGGGAAC 238
360 CACCACCTAAAGTCGGAAATCTACCTGATCAGAGAGAGAGATGATCAGGGTCCAG 419
239 CAGCAACTCAACGTCAGATCCTCAGCTGCTCAGGAGGGAGAGATG--AGGGAGCAT 295
420 CTGAGATTCAAGTGCCTCAGCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTG 479
296 CTGAGGTCAGAGGCCGAGGCTGAGCTGATGATGAGCAGGAGACAGGGTCAACCCAGACTG 355
480 GGGATGATGTAAGTGGTACTGATGATCAAGGGGAGAGATCTACCAAAAGCAGAGCACT 539
356 GGTGTGAGTGAAGATGGTCTGATGGGAGGAGATGGACCCGCCAAATCCAGAGGAGG 415
540 TTAATAATCCAGAGCAGTGAAGGGAATCAGAGGTTTAAAGGAGATGATGATGAAACA 599
416 TGAACAACGCTGAGAGAGGTGAAGCAATCAGCTGTTTAAAGAGGCACGTTGAATG 475
600 ACACAACTGTTTATATAGATATTTTACTTTAAAGAGTCT 642
476 ATGCAAGGCTGCTCTATGTTGAAATTTGTTCAATTAATTTCT 518

RESULT 32

US-08-669-161A-15
; Sequence 15, Application US/08669161A
; Patent No. 6013481
; GENERAL INFORMATION:
; APPLICANT: DeBacker, Olivier; Van den Eynde,
; APPLICANT: Benoit; Boon-Falleur, Thierry
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,161A
; FILING DATE: 24-June-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/531,662
; FILING DATE: 21-September-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-January-1995
; APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-May-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-July-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6013481man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5443
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 560 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-669-161A-15
Query Match 24.2%; Score 162.2; DB 3; Length 560;
Best Local Similarity 62.4%; Pred. No. 4,9e-42;
Matches 289; Conservative 0; Mismatches 168; Indels 6; Gaps 2;
QY 180 ATAGGTCCCGAGGCGAGTCTCCAGGAGCACTGAATAATGTAAGTGGCCGAGAA 239
Db 62 ATAGGTGATCTCTCTCGCCAACTCATATTTTACACAGATGAATCTCAGTAGAGAAAAT 121
QY 240 GATCAACATATAGGCTTAGGCCCAAGAAAGTTTACAGCTCTCCAGCTGATTGGGGCTA 299
Db 122 CGAOCCTATTATTGGCTTAGACCAAGCGCTATGTACAGCTCTCTGAGTGTGGSCCTA 181
QY 300 TGTCTACTGGCTCCCTTTGTCCAGGAGCACTGATGAAGAGCTTAAAGAGAGAAAC 359
Db 182 TCGGCCCCGAG--CAGTTTCAGTGATGAAGTGGACACGACACCTTGAGAGGGGAAC 238
QY 360 CACCACCTAAAGTGGAAATCTACCTGATCAGAGAGAGAGATGATCAGGGTGGAG 419
Db 239 CAGCAACTCAACGTCAGATCCTCAGCTGCTCAGGAGGAGAGATG---AGGAGCAT 295
QY 420 CTGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTG 479
Db 296 CTGAGGTCAGAGGCCGAGGCTGAGCTGATGAGCAGGAGACAGGGTCAACCCAGACTG 355
QY 480 GGGATGATGTAAGTGGTACTGATGTCAGGGGAGAGATCTACCAAAAGCAGAGCACT 539
Db 356 GGTGTGAGTGAAGATGGTCTGATGGGAGGAGATGGACCCGCCAAATCCAGAGGAGG 415
QY 540 TTAATAATCCAGAGCAGTGAAGGGAATCAGAGGTTTAAAGGAGATGATGATGAAACA 599
Db 416 TGAACAACGCTGAGAGAGGTGAAGCAATCAGCTGTTTAAAGAGGCACGTTGAATG 475
QY 600 ACACAACTGTTTATATAGATATTTTACTTTAAAGAGTCT 642
Db 476 ATGCAAGGCTGCTCTATGTTGAAATTTGTTCAATTAATTTCT 518

RESULT 33

US-08-602-039-15
; Sequence 15, Application US/08602039
; Patent No. 6069001
; GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry;
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,039
; FILING DATE: 15-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/370,648
; FILING DATE: 10-JANUARY-1995
; APPLICATION NUMBER: 08/250,162

FILING DATE: 27-MAY-1994
PRIOR APPLICATION NUMBER: 08/096,039
APPLICATION DATE: 22-JULY-1993
FILING DATE: 22-JULY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6069001man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5323.2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 560 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-502-039-15

Query Match 24.2%; Score 162.2; DB 3; Length 560;
Best Local Similarity 62.4%; Pred. No. 4.9e-42;
Matches 289; Conservative 0; Mismatches 168; Indels 6; Gaps 2;

QY 180 ATAGGTCCGAGGCGAGTCTCCAGGAAGTGAATAGTGAATATGAGTTGGCGAGGAA 239
DB 62 ATAGTGTGATCTTCTCGCCCACTCATATTTACACAGATGAATCTCAGTAGAGGAAAT 121
QY 240 GATCAACATATAGCGCTAGGCCCAAGAGAGTTTACAGCCCTCTAGCTGATTTGGGGCTA 299
DB 122 CGACCTATTATTGGCCCTAGACCAAGGCGCTATGTACAGCCCTCTAGTGAATTTGGCCCTA 181
QY 300 TGCTTACTGTGCTCCCTTTGTCGCCAGGACCCACTGATGAAGAGCTTAAGAGAGGAAAC 359
DB 182 TGGGCCCCGAG--CAGTTTCAAGTATGATGAGTGAACACAGTGAAGAGGGAAC 338
QY 360 CACCACATAAAGTCGGAATCTTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAG 419
DB 239 CAGCAACTCAAGCTCAGGATCTCTGAGTCTGCTCAGGAGGAGAGGATG---AGGGAGCAT 295
QY 420 CTGAGATTCAAGTCCCTGACCTGGAAGCCGATCTCCAGGAGCTATCTCAGACAAAGCATG 479
DB 296 CTGAGTCAAGGCCGAGGCTGAGCTGATGATGAGGAGGAGGCTACCCACAGACTG 355
QY 480 GGGATGATGTGAAGTGGTACTGATGATCAAGGGGAGATTTTACCAAAGCAGAGCACT 539
DB 356 GGTGTGATGTGAAGATGCTCTGATGGCAGGAGATGGACCCGCCCAATCCAGAGGAG 415
QY 540 TTAATGTCAGAGAGAGTGAAGGGAATCAAGGTTTAAAGAGAGATAGCTGAACA 599
DB 416 TGAAGAGCCCTGAAGAGGTTGAAGCAATCAGAGTGTAAAGAGGAGGACGTTGAATG 475
QY 600 ACACAACTGTTTTATATAGATATTTTACTTTAAAGAGTCT 642
DB 476 ATGAGGCTGCTCTATGTTGGAAATTTGTTCAATTAATTTCT 518

RESULT 34

US-09-702-705-1827
Sequence 1827, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C14

CURRENT APPLICATION NUMBER: US/09702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 1827
LENGTH: 288
TYPE: DNA
ORGANISM: Homo sapiens
US-09-702-705-1827

Query Match 23.2%; Score 155.2; DB 4; Length 288;
Best Local Similarity 77.3%; Pred. No. 5.9e-40;
Matches 201; Conservative 0; Mismatches 58; Indels 1; Gaps 1;

QY 319 GTCCAGGAACCCACTGATGAAGAGCCTTAAAGAGAGAAACCCACCTAAAGTCCGAA 378
DB 29 GTCCAGGAGCCAGTATGAGAGGCCCAAAAGAGAGACAGCTGAAGTCCGGA 88
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGA 438
DB 89 TCCTACACCTGGCAGCAGACAGAGAGATCAGGATACAGCTGAGATCCCGAGTCCGGA 148
QY 439 CTTGAAGCCGATCTCCAGGAGCTATGTACAGCAAAAGACTGGGATGGATGTGAAGT-G 497
DB 149 CATGGAAGTGTATCTGCAAGAGCTGCATCACTCAACACCGGGGATTAATCTGATTTGG 208
QY 498 GTACTGATGTCAAGGGGAGATTTTACCAAAGCAGAGCACTTTAAATGCCAGAGCAG 557
DB 209 GTTCCGGGCTCAAGGTGAAGATAATACCTAAAGAGGAACACTGTAAATGCCAGAGCAG 268
QY 558 GTGAAGGGAAATCACAGTT 577
DB 269 GTGAAGAGCAACCAAGTT 288

RESULT 35

US-09-736-457-1827
Sequence 1827, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 1827
LENGTH: 288
TYPE: DNA
ORGANISM: Homo sapiens
US-09-736-457-1827

Query Match 23.2%; Score 155.2; DB 4; Length 288;
Best Local Similarity 77.3%; Pred. No. 5.9e-40;
Matches 201; Conservative 0; Mismatches 58; Indels 1; Gaps 1;

QY 319 GTCCAGGAACCCACTGATGAAGAGCCTTAAAGAGAGAAACCCACCTAAAGTCCGAA 378
DB 29 GTCCAGGAGCCAGTATGAGAGGCCCAAAAGAGAGACAGCTGAAGTCCGGA 88
QY 379 TCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGA 438
DB 89 TCCTACACCTGGCAGCAGACAGAGAGATCAGGATACAGCTGAGATCCCGAGTCCGGA 148
QY 439 CTTGAAGCCGATCTCCAGGAGCTATGTACAGCAAAAGACTGGGATGGATGTGAAGT-G 497
DB 149 CATGGAAGTGTATCTGCAAGAGCTGCATCACTCAACACCGGGGATTAATCTGATTTGG 208
QY 498 GTACTGATGTCAAGGGGAGATTTTACCAAAGCAGAGCACTTTAAATGCCAGAGCAG 557
DB 209 GTTCCGGGCTCAAGGTGAAGATAATACCTAAAGAGGAACACTGTAAATGCCAGAGCAG 268
QY 558 GTGAAGGGAAATCACAGTT 577
DB 269 GTGAAGAGCAACCAAGTT 288

Db 89 TCCTACACCTGGCAGCAGACAGAGAGATCAGGATACAGCTGAGATCCAGTCCAGTGGCGA 148
Qy 439 CTTGGAAGCCGATCTCCAGAGCTATGTGACAGAAACAGCTGGGATGGATGTGAAGGT-G 497
Db 149 CATGGAAGGTGATCTGCAAGAGCTGATCAGTCAACACGGGGATAAATCTGGATTTGG 208
Qy 498 GTACTGATGTCAGGGGGAAGATCTTACCAAAGCAGAGCACTTTAAATGCCAGAGACAG 557
Db 209 GTTCCGGCGTCAGGTGAAGATATACCTTAAGAGGACACTGTAAATGCCAGAGACG 268
Qy 558 GTGAAGGGAATCACAGTT 577
Db 269 GTGAAGAGCAACCAAGTT 288

RESULT 36
US-09-328-475C-263
; Sequence 263, Application US/09328475C
; Patent No. 6476207
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jimmy
; APPLICANT: Astel, Jon H.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Endege, Wilson O.
; APPLICANT: Ford, Donna M.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; APPLICANT: Steinmann, Kathleen E.
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT
; TITLE OF INVENTION: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER
; FILE REFERENCE: 1532.002/200130.463
; CURRENT APPLICATION NUMBER: US/09/328.475C
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 263
; LENGTH: 656
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(656)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-475C-263

Query Match 22.2%; Score 148.6; DB 4; Length 656;
Best Local Similarity 62.0%; Pred. No. 1.2e-37;
Matches 291; Conservative 0; Mismatches 136; Indels 42; Gaps 2;

Qy 90 TGCCACCTGGTGACCGCTGGGGAGCTGTGAGAGTGTGAGGGGACGTTCCAGCCGCTG 149
Db 82 TACCGCGGGGAGAACGCCAGGAGCTGTGAGAGTGTGAGTGGCTTCTGCTGTCGG 141
Qy 150 GACTCTTTCTCTCTCTGAGAGCGGCTATAGTTCGGAGGCGAGTCTCCAGGAC 209
Db 142 ACACATTTTCTCTCTCTGAGACTCATCTGGTAGATCCGAGGCGAGTCTCCAGGGGC 201
Qy 210 TCAAAATAGTGAATATGATGTGGCAGGAAGATCAACATATAGGCTAGGCCAAGAAGAA 269
Db 202 TGAAGTTGGAATATGGTTTCTAAGAGATTAATCTAICGGGGTAGACCAATGATCT 261
Qy 270 GTTTAAGCCTCTGAGCTGATTTGGGCTATGCTTACTGCTTCCCTTTGTTCCAGGAC 329
Db 262 ATGTAGAATCTTCTG-----AGGAGT 282
Qy 330 CCAGTATCAAGAGCCTTAAGAGAGAAACCCACCTAAAGTGGGAATCTACACCTG 389
Db 283 CAGTATGAGCAACCTGAGAGTGAATCAACCACTCAAGTCAAGTCAAGTCTACACCTG 342
Qy 390 ATCAGAAGAGAGAGATGATCAGGGTGCAGGTTCAAGTTCAGCTGCTGAGGAGCGG 449
Db 343 CTGAAGAGAGAGAGGATG---AGGAGCATCTGCAGCTCAAGGGCAGGAGCCTGAAGCTG 399

Qy 450 ATCTCCAGGAGCTATGTACAGCAAGACTGGGATGGAGTGAAGTGGTGGTACTGATCTCA 509
Db 400 ATAGCCAGGAGCTGTTTACGCAAGAGACTGGTGTGAGCTTGGAGATGGTCTGTGATACCA 459
Qy 510 AGGGGAAGATCTTACCAAAAGCAGGACACTTTAAATGCCAGAGACGAGG 558
Db 460 AAGGGGTTGCTCGGAAATGAAGAGCAGATGAACCTGCCCGNAGAGG 508

RESULT 37
US-08-370-648-1
; Sequence 1, Application US/08370648
; Patent No. 5648226
; GENERAL INFORMATION:
; APPLICANT: Van den Eynde, Benoit; DeBacker, Olivier;
; APPLICANT: Boon-Falleur, Thierry;
; TITLE OF INVENTION: Isolated, Truncated Nucleic Acid
; TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,
; TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,648
; FILING DATE: 10-JANUARY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,162
; FILING DATE: 27-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/096,039
; FILING DATE: 22-JULY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5648226man D.
; REGISTRATION/DOCKET NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5323.2
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 646 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-370-648-1

Query Match 21.6%; Score 144.2; DB 1; Length 646;
Best Local Similarity 68.0%; Pred. No. 3.1e-36;
Matches 232; Conservative 0; Mismatches 103; Indels 6; Gaps 2;

Qy 217 GTCAATATGAGTTGGCGAGGAGATCAACATATAGGCTAGGCCAAGAAGATTACA 276
Db 42 GTGAATATGAGTTGGCGAGGAGATCGACCTATCGGCTAGACCAAGACGCTACGTAGA 101
Qy 277 GCCTCTGAGCTGATTTGGGCTATGCTTATCGGCTCCCTTTGTCCAGGAACCCACTGA 336
Db 102 GCCTCTGAAATGATTTGGGCTATCGGCTATCGGCTAGACCAAGACGTTGAAGTGAAC 158
Qy 337 TGAGAGCCTTAAGAGAGAGAAACCCACTAAAGTTCGAAATCCTACACCTGATCAGAA 396
Db 159 AGCAACACCTTGAAGAGAGGAGAACCAAGAACTCAAGCTCAGGATCTCTGAGCTGCTCAGGA 218

TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 646 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-602-039-1

Query Match 21.6%; Score 144.2; DS 3; Length 646;
Best Local Similarity 68.0%; Pred. No. 3.1e-36;
Matches 232; Conservative 0; Mismatches 103; Indels 6; Gaps 2;
QY 217 GTGAATAATGAGTTGGCGAGAGAGATCAACATATAGCGCTAGGCCAAGAGAGTTTACA 276
Db 42 GTGAATAATGAGTTGGCGAGAGAGATCGACTATCGGCCCTAGACCAAGACGTCAGTAGA 101
QY 277 GCCTCCTGAGCTGATTCGGGGCTATGCTTACTTGGCTCCCTTTGTCCAGGAACCACTGA 336
Db 102 GCCTCCTGAAATGATTTGGGCTATGCGCC---CGAGCAGTTTCAGTGATGAAGTGAACC 158
QY 337 TGAAGAGCCTTAAGAAGAGAGAAACCCACCTATAAAGTCGGATCTCTACCTGTATCAGAA 396
Db 159 AGCAACACCTGAAGAAGGGGAAACCAAGCACTCAACGTCAGGATCTCTGCACTGCTCAGGA 218
QY 397 GAGAGAAGATGATCAGGGTGCAGCTGAGATTCAAGTGCGCTGACCTGGAAGCCGATCTCCA 456
Db 219 GGGAGAGATG---AGGAGCATCTTCAGGTCAAGGGCCGAGCCTGAAGCTGATAGCCA 275
QY 457 GGAGCTATGTCAGACAAAGACTGGGGATGAGATGTGAAGTGGTGTATGTATGTCAAGGGGA 516
Db 276 GGAACAGGGTCAACCCACAGACTGGGTGTGAGTGTGAAGATGCTCTGATGGCGCAGGAGAT 335
QY 517 GATTCTACCAAAAGCAGACGACTTTAAATATGCCAGAAAGCAG 557
Db 336 GGACCGCCAAATCCAGAGAGAGGTGAAAACGCTGAGAGAG 376

RESULT 41
US-08-250-162A-1
Sequence 1, Application US/08250162A
Patent No. 5610013
GENERAL INFORMATION:
APPLICANT: Van den Eynde, Benoit; Boon-Falleur, Thierry
TITLE OF INVENTION: METHOD FOR DIAGNOSING A DISORDER BY
TITLE OF INVENTION: DETERMINING EXPRESSION OF GAGE TUMOR REJECTION ANTIGEN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESSES:
ADDRESSEES: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/250,162A
FILING DATE: 27-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/096,039
FILING DATE: 22-JULY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5610013man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5323.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884


```
Qy 453 TCAGGAGCTATGTGACACAAAGACTGGGATGGATGTGAAGTGGTACTGATGTCAAGG 512
Db 344 GCCAGGAACAGGGTCAACCCACAGACTGGTGTGAGTGTGAAGTGGTCTGATGGCAGG 285
Qy 513 GGRAGATTCACAAAGCAGACACTTAAATGCCAGAGCAGGTGAAGGGARATCAC 572
Db 284 AGATGGACCCGCCAATCCAGAGAGGTGAAGCGCTGAAGAGGTGAAGCAATCAC 225
Qy 573 AGGTTAAAGAGATAGCTGAAACACAACTGTTTTATATTAGATATTTTACT 632
Db 224 AGTGTTAAAGAGGCGGTTGAATGATGAGCTGCTCTATGTTGGAAATTTGTCA 165
Qy 633 TAAGAGCTCT 642
Db 164 TTAAGATCT 155
```

RESULT 44

US-09-439-313-472

; Sequence 472, Application US/09439313

; Patent No. 6329505

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan Louise

; APPLICANT: Jiang Yuqui

; APPLICANT: Reed, Steven G.

; APPLICANT: Kalos, Michael

; APPLICANT: Fanger, Gary

; APPLICANT: Retter, Mark

; APPLICANT: Solk, John

; APPLICANT: Day, Craig

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND

; FILE REFERENCE: 210121.427C9

; CURRENT APPLICATION NUMBER: US/09/439,313

; NUMBER OF SEQ ID NOS: 575

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 472

; LENGTH: 515

; TYPE: DNA

; ORGANISM: Homo sapiens

; NAME/KEY: misc feature

; LOCATION: (1)...(515)

; OTHER INFORMATION: n = A,T,C or G

US-09-439-313-472

```
Query Match 14.8%; Score 99.2; DB 4; Length 515;
Best Local Similarity 63.0%; Pred. No. 6.8e-22;
Matches 174; Conservative 0; Mismatches 93; Indels 9; Gaps 1;
```

```
Qy 341 GAGCCTAAAGAGAGAAACCCCACTAAAGTCGGAATCCTACACCTGATCAGAAGAGA 400
Db 152 GAATCTCAGCAGAGAGAACCCCACTGACATCAGGATATTGAACCTGGACAGAGAGA 211
Qy 401 GAAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAAGCCGATCTCCAGGAG 460
Db 212 GAA-----GGAACACCTCCGATCGAAGAACCTAAAGTAGAAGGTGATTGCCAGGAA 262
Qy 461 CTATGTCAGACAAAGACTGGGGATGGATGTGAAGTGGTACTGTCAAGGGGAAGATT 520
Db 263 ATGGATCTGGAAAGACTCGGAGTGGAGTGGATGCTCTGATGTAAAGAGAGACT 322
Qy 521 CTACCAAAAGCAGACACTTTTAAATGCCAAGCAGGTGAAGGGAATCAGGTTTAA 580
Db 323 CCACCTAATCCTAAGCATGTAAAGACTAAAGAGCAGGAGATGGCAGCCATAAGTTAAA 382
Qy 581 AGGAGATAGCTGAAACACACAACTGTTTTAT 616
Db 383 AAGAAGACAGCTGAAGCTACACATGGCTGATGT 418
```

RESULT 45

US-09-352-616A-472

; Sequence 472, Application US/09352616A

; Patent No. 6395278

; GENERAL INFORMATION:

; APPLICANT: Dillon, Davin C.

; APPLICANT: Harlocker, Susan Louise

; APPLICANT: Jiang, Yuqui

; APPLICANT: Xu, Jiangchun

; APPLICANT: Mitcham, Jennifer Lynn

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS

; FILE REFERENCE: 210121.427C8

; CURRENT APPLICATION NUMBER: US/09/352,616A

; CURRENT FILING DATE: 1999-07-13

; NUMBER OF SEQ ID NOS: 472

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 472

; LENGTH: 515

; TYPE: DNA

; ORGANISM: Homo sapiens

; NAME/KEY: misc feature

; LOCATION: (1)...(515)

; OTHER INFORMATION: n = A,T,C or G

US-09-352-616A-472

```
Query Match 14.8%; Score 99.2; DB 4; Length 515;
Best Local Similarity 63.0%; Pred. No. 6.8e-22;
Matches 174; Conservative 0; Mismatches 93; Indels 9; Gaps 1;

Qy 341 GAGCCTAAAGAGAGAAACCCCACTAAAGTCGGAATCCTACACCTGATCAGAAGAGA 400
Db 152 GAATCTCAGCAGAGAGAACCCCACTGACATCAGGATATTGAACCTGGACAGAGAGA 211
Qy 401 GAAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAAGCCGATCTCCAGGAG 460
Db 212 GAA-----GGAACACCTCCGATCGAAGAACCTAAAGTAGAAGGTGATTGCCAGGAA 262
Qy 461 CTATGTCAGACAAAGACTGGGGATGGATGTGAAGTGGTACTGTCAAGGGGAAGATT 520
Db 263 ATGGATCTGGAAAGACTCGGAGTGGAGTGGATGCTCTGATGTAAAGAGAGACT 322
Qy 521 CTACCAAAAGCAGACACTTTTAAATGCCAAGCAGGTGAAGGGAATCAGGTTTAA 580
Db 323 CCACCTAATCCTAAGCATGTAAAGACTAAAGAGCAGGAGATGGCAGCCATAAGTTAAA 382
Qy 581 AGGAGATAGCTGAAACACACAACTGTTTTAT 616
Db 383 AAGAAGACAGCTGAAGCTACACATGGCTGATGT 418
```

Search completed: January 18, 2004, 18:48:33

Job time : 40.4394 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 23:43:50 ; Search time 1008.55 Seconds
(without alignments)
16121.802 Million cell updates/sec

Title: US-10-051-835-1
Perfect score: 669
Sequence: 1 cacaacgcagccagcactt.....tttttggcagtcgatctc 669

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

EST:

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hic:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_est3:*
- 12: gb_est4:*
- 13: gb_est5:*
- 14: gb_est6:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_dhg:*
- 27: em_gss_vri:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	566.6	84.7	673	14	CB995986
2	558	83.4	756	10	BE903565
3	549.8	82.2	706	10	BE281431
4	545.6	81.6	780	10	BE734431

5	540	80.7	652	10	BG623213
6	538.4	80.5	592	10	AW972530
7	536.8	80.2	908	10	BG336433
8	531	79.4	782	10	BG336384
9	506.6	75.7	588	10	BE733225
10	496.8	74.3	982	10	BE900651
11	484.4	69.4	438	13	BX283754
12	445.2	66.5	533	14	T87105
13	442.6	66.2	496	9	AA993199
14	438.6	65.6	496	9	AI188535
15	438.4	65.5	543	14	N30289
16	426.6	63.8	483	9	AI149003
17	404.6	60.5	468	9	AI128458
18	404	60.4	468	9	AA039331
19	391.6	58.5	453	9	AI189715
20	377.6	56.4	532	9	AA737311
21	376.4	56.3	493	14	H53211
22	372.6	55.7	422	9	AA505794
23	372.2	55.6	512	14	W93043
24	364.6	54.5	445	10	BE302172
25	360.6	53.9	421	9	AA722699
26	357.4	53.4	437	9	AI740470
27	351.8	52.6	555	14	T87104
28	343.8	51.4	429	9	AI200158
29	334.6	50.0	404	14	H95793
30	318	47.5	349	9	AI186558
31	316.8	47.4	401	14	H53104
32	316.2	47.3	389	14	W93171
33	313.4	46.8	420	14	N42010
34	304.8	45.6	547	14	N40147
35	303.2	45.3	410	14	R99925
36	296.4	44.3	570	14	CB992682
37	296.4	44.3	570	14	CB994958
38	290.2	43.4	499	10	BG354572
39	286.4	42.8	520	14	CB961487
40	285.2	42.6	524	11	AF318372
41	280.8	42.0	503	9	AI745551
42	278.6	41.6	849	10	BG436305
43	277.8	41.5	429	9	AI189067
44	266.8	39.9	544	14	H87913
45	266.4	39.8	430	14	N27023
46	245.2	36.7	402	14	H87367
47	240.6	36.0	534	10	BG436468
48	238.4	35.6	571	14	N32547
49	234.8	35.1	347	9	AA039416
50	232.8	34.8	374	14	H96141
51	229	34.2	394	14	D59091
52	215.2	32.2	387	14	T84951
53	214.4	32.0	382	13	BX283272
54	204	30.5	414	9	AA620697
55	204	30.5	428	9	AA948244
56	204	30.5	438	9	AI954712
57	204	30.5	441	13	BX101350
58	204	30.5	486	12	BI460303
59	204	30.5	487	12	BG772292
60	204	30.5	489	10	BE881023
61	204	30.5	491	13	BU182421
62	204	30.5	499	13	BU564534
63	204	30.5	880	10	BE876234
64	204	30.5	1140	10	BE620196
65	202.4	30.3	421	9	AI631681
66	202.4	30.3	423	9	AA648117
67	201	30.0	771	13	BU929434
68	200.8	30.0	394	10	BG216048
69	200.8	30.0	485	10	BG217244
70	194.6	29.1	556	10	AW975988
71	192.2	28.7	978	13	BU532066
72	191.4	28.6	407	9	AA758618
73	188.2	28.1	487	10	BG211375
74	184.6	27.6	388	10	BG206865
75	183.4	27.4	391	10	BG205627
76	182	27.2	271	14	N41436
77	181.2	27.1	569	10	BG700165

```

78 179.8 26.9 1248 10 BG191783
79 174.8 26.1 407 10 BG193320
80 174.8 26.1 409 10 BG221249
81 173.2 25.9 406 10 BG208252
82 172.2 25.7 389 10 BG205030
83 171.4 25.5 521 13 EX108227
84 170.8 25.5 521 12 BG773070
85 169.8 25.4 537 12 BM82793
86 169.2 25.3 669 10 BG435916
87 168.4 25.2 505 14 CB115693
88 168.2 25.1 457 9 AW510753
89 168.2 25.1 527 12 B1826605
90 166.6 24.9 509 12 B1868671
91 164.4 24.6 464 9 AA738037
92 164.2 24.5 572 10 BG120838
93 163.4 24.4 905 10 BF239019
94 163.2 24.4 340 10 BE620755
95 161.4 24.1 665 9 AV682444
96 161 24.1 450 9 AA760996
97 161 24.1 468 9 AI392665
98 160.8 24.0 851 10 BG250953
99 159.6 23.9 221 9 AA039250
100 158.8 23.7 447 9 AI381509

```

ALIGNMENTS

```

RESULT 1
LOCUS CB995986 673 bp mRNA linear EST 01-MAY-2003
DEFINITION AGENCOURT 13511631 NIH_MGC_148 Homo sapiens cDNA clone
IMAGE:30330532.5', mRNA sequence.

```

```

ACCESSION CB995986
VERSION CB995986.1
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)

```

```

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgs.nci.nih.gov/
JOURNAL 1 (bases 1 to 673)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM344 row: i column: 05
High quality sequence stop: 584.

```

FEATURES

```

source
1..673
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30330532"
/tissue="pre-eclampsia placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_148"
/notes="Organ: placenta; Vector: pBluescriptR; Site: 1:
all-XhoI; Site 2: BamHI; Library is oligo-dT primed and
directionally cloned using primer
5'-TTTTTTTTTTTTTTT-3', size-selected for average insert
size 2.3 kb and normalized to 10^5. This is a primary
library, enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation).
Library constructed by M. Brownstein (NHGRI/NHGR1,

```

```

National Institutes of Health). Note: this is a NIH_MGC
Library."
BASE COUNT 205 a 145 c 169 g 152 t 2 Others
ORIGIN
Query Match 84.7%; Score 566.6; DB 14; Length 673;
Best Local Similarity 95.8%; Pred. No. 5.5e-138;
Matches 600; Conservative 0; Mismatches 5; Indels 21; Gaps 1;
QY 29 ATGTTCTTTGGACACCTCGCTCAGTGTGTCATGTTCACTGGGCACTTCCCTTCGACCCCT 88
DB 28 ATGTTCTTTGGACACCTCGCTCAGTGTGTCATGTTCACTGGGCACTTCCCTTCGACCCCT 87
QY 89 TTGCCACGTTGGTACCCCTCGGAGCTGTGAGAGTGTGAGGGGACAGTTCACCGCTCT 148
DB 88 TTGCCACGTTGGTACCCCTCGGAGCTGTGAGAGTGTGAGGGGACAGTTCACCGCTCT 147
QY 149 GGACTCTTTCTCTCTACTGAGAGCGAGCTATAGGTCGCGAGGCGAGTCTCCAGGAA 208
DB 148 GGACTCTTTCTCTCTACTGAGAGCGAGCTATAGGTCGCGAGGCGAGTCTCCAGGAA 207
QY 209 CTGAATAGTGAATATGAGTTGGCGAGGAGATCAACATATAGGCTTAGGCCCAAGAGA 268
DB 208 CTGAATAGTGAATATGAGTTGGCGAGGAGATCAACATATAGGCTTAGGCCCAAGAGA 267
QY 269 AGTTTACAGCTCTCTGAGCTGATTTGGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAA 328
DB 268 AGTTTACAGCTCTCTGAGCTGATTTGGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAA 306
QY 329 CCCACTGATGAAGAGCTCTAAAGAGAGAGAAACACCCACTAAAAGTCGGAATCTTACACCT 388
DB 307 CCCACTGATGAAGAGCTCTAAAGAGAGAGAAACACCCACTAAAAGTCGGAATCTTACACCT 366
QY 389 GATCAGAGAGAGAGAGATGATCAGGTCGAGCTGAGATTCAGTGCCTGACCTGGAAGCC 448
DB 367 GATCAGAGAGAGAGAGATGATCAGGTCGAGCTGAGATTCAGTGCCTGACCTGGAAGCC 426
QY 449 GATCTCCAGGAGCTATGTCAGACAAAGAGCTGGGATGATGTGAAGTGGTACTGATGTC 508
DB 427 GATCTCCAGGAGCTATGTCAGACAAAGAGCTGGGATGATGTGAAGTGGTACTGATGTC 486
QY 509 AAGGGAGAGATCTTACCAAAGAGAGACACTTTTAAATGCCAGAGCAGGTGAAGGGAAA 568
DB 487 AAGGGAGAGATCTTACCAAAGAGAGACACTTTTAAATGCCAGAGCAGGTGAAGGGAAA 546
QY 569 TCACAGGTTTAAAGAGAGATGAAGCTGAAACAAACAACTGTTTTTATATTAGATATTTT 628
DB 547 TCACAGGTTTAAAGAGAGATGAAGCTGAAACAAACAACTGTTTTTATATTAGATATTTT 606
QY 629 ACTTTAAAGAGCTCTTAATAATTTT 654
DB 607 ACTTTAAATATCTTAATAAGTTTT 632

```

```

RESULT 2
LOCUS BE903565 756 bp mRNA linear EST 29-SEP-2000
DEFINITION 601677770F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3960352.5',
mRNA sequence.
ACCESSION BE903565
VERSION BE903565.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgs.nci.nih.gov/
JOURNAL 1 (bases 1 to 756)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: ATCC

```

cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCM842 row: 1 column: 17
 High quality sequence start: 34
 High quality sequence stop: 705.
 Location/Qualifiers

FEATURES

1..756
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3960352"
 /tissue_type="choriocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC_21"
 /note="Organ: placenta; Vector: pOTB7; Site: 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 223 a 150 c 216 g 167 t

ORIGIN

Query Match 83.4%; Score 558; DB 10; Length 756;
 Best Local Similarity 95.8%; Pred. No. 1e-135; 5; Indels 21; Gaps 1;
 Matches 592; Conservative 0; Mismatches 5;
 37 TTGACACCTGCCTCAGTGTGCATGTTCACTGGGCATCTTCCCTCGACCCCTTTGCCAC 96
 41 TTGACACCTGCCTCAGTGTGCATGTTCACTGGGCATCTTCCCTCGACCCCTTTGCCAC 100
 97 GTGGTACCGCTGGGAGCTGTGAGAGTGTGAGGGCAGTTCCAGCCGTCTGGACTCTT 156
 101 GTGGTACCGCTGGGAGCTGTGAGAGTGTGAGGGCAGTTCCAGCCGTCTGGACTCTT 160
 157 TCTCTCTACTGAGAGCGAGCTATAGTTCGGAGCCAGTCTCTCCAGGAACCTGAATA 216
 161 TCTCTCTACTGAGAGCGAGCTATAGTTCGGAGCCAGTCTCTCCAGGAACCTGAATA 220
 217 GTGAATATAGTGTGGCGAGGAAGATCAACATATAGGCTAGCCAGAGGAAGTTTACA 276
 221 GTGAATATAGTGTGGCGAGGAAGATCAACATATAGGCTAGCCAGAGGAAGTTTACA 280
 277 GCCTCTGAGCTGATTGGGGCTATGCTTACTGGCTCCCTTTGTCGAGGAACCCACTGA 336
 281 GCCTCTGAGCTGATTGGGGCTATGCTT-----GACCCACTGA 319
 337 TGAAGAGCCTTAAGAGAGAGAAACCCCACTAAAGTCGGAATCTTACACCTGATCAGAA 396
 320 TGAAGAGCCTTAAGAGAGAGAAACCCCACTAAAGTCGGAATCTTACACCTGATCAGAA 379
 397 GAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAGCGATCTCCA 456
 380 GAGAGAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAGCGATCTCCA 439
 457 GGAGCTATGTGAGCAACAAGACTGGGGATGGATGTGAAGTGTGATGTCAAGGGGAA 516
 440 GGAGCTATGTGAGCAACAAGACTGGGGATGGATGTGAAGTGTGATGTCAAGGGGAA 499
 517 GATTCTACCAAGAGCAGAGCACTTAAATGTCAGAGCAGGTGAGGGAATTCACAGGT 576
 500 GATTCTACCAAGAGCAGAGCACTTAAATGTCAGAGCAGGTGAGGGAATTCACAGGT 559
 577 TTAAGGAAGATAGCTGAAACCAACCAAACTCTTTTATATATAGATATTTTACTTTAAA 636
 560 TTAAGGAAGATAGCTGAAACCAACCAAACTCTTTTATATATAGATATTTTACTTTAAA 619
 637 GAGTCTTAATAAATTTT 654

||||| 620 ATATCTAATAAAGTTT 637
 BE281431 706 bp mRNA linear EST 13-JUL-2000
 601154945F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3138515 5',
 mRNA sequence.
 BE281431
 BE281431 GI:9156450
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 706)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: rga@bbs-k@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCM103 row: 1 column: 12
 High quality sequence stop: 628.
 Location/Qualifiers
 1..706
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3138515"
 /tissue_type="choriocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_21"
 /note="Organ: placenta; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 216 a 158 c 185 g 147 t

Query Match 82.2%; Score 549.8; DB 10; Length 706;
 Best Local Similarity 95.4%; Pred. No. 1.5e-133;
 Matches 596; Conservative 0; Mismatches 7; Indels 22; Gaps 2;
 36 TTGACACCTGCCTCAGTGTGCATGTTCACTGGGCATCTTCCCTCGACCCCTTTGCCCA 95
 1 TTGACACCTGCCTCAGTGTGCATGTTCACTGGGCATCTTCCCTCGACCCCTTTGCCCA 60
 96 CGTGGTGCACCGCTGGGAGCTGTGAGAGTGTGAGGGCAGTTCACGCGCTCTGGACTCT 155
 61 CGTGGTGCACCGCTGGGAGCTGTGAGAGTGTGAGGGCAGTTCACGCGCTCTGGACTCT 120
 156 TTCTCTCTACTGAGACGCGACCTTATAGTTCGAGCCAGTCTCTCCAGGAACCTGAAT 215
 121 TTCTCTCTACTGAGACGCGACCTTATAGTTCGAGCCAGTCTCTCCAGGAACCTGAAT 180
 216 AGTGAATATAGTTCGAGGAGAGATCAACATATAGGCTAGCCAGGAGAGATTAC 275
 181 AGTGAATATAGTTCGAGGAGAGATCAACATATAGGCTAGCCAGGAGAGATTAC 240
 276 AGCTCTCTGAGCTGATTGGGGCTATGCTTACTGGCTCCCTTTGTCGAGGAACCCACTG 335
 241 AGCTCTCTGAGCTGATTGGGGCTATGCTT-----GACCCACTG 279

```

336 ATGAAGAGCCTAAAGAGAGAAACCCACCACTAAAGTCGGAATCCTACACCTGATCAGA 395
280 ATGAAGAGCCTAAAGAGAGAAACCCACCACTAAAGTCGGAATCCTACACCTGATCAGA 339
396 AGAGAGAGATGATCAGGTCGACGTGAGATTCAGTGCCTGACCTGGAAGCGGATCTCC 455
340 AGAGAGAGATGATCAGGTCGACGTGAGATTCAGTGCCTGGAAGCGGATCTCC 399
456 AGAGAGATGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 515
400 AGAGAGATGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 459
516 AGATTCACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 575
460 AGATTCACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 519
576 TTTA-AAGGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 634
520 TTACAGGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 579
635 AAGAGCTCTTAATAATTTTGGCAT 659
580 AATATCTTAATAAGTTTAAGCTT 604

RESULT 4
BE734431
LOCUS 60156206F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3841272 5',
DEFINITION mRNA sequence.
ACCESSION BE734431
VERSION BE734431.1 GI:10148423
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 780)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LCM532 row: k column: 01
High quality sequence stop: 652.
FEATURES
Location/Qualifiers
1..780
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3841272"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_21"
/notes="Organ: placenta; Vector: pOTB7; Site:1: XhoI;
Site:2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 247 a
ORIGIN 148 c 215 g 170 t

Query Match 81.6%; Score 545.6; DB 10; Length 780;

```

Best Local Similarity 95.9%; Pred. No. 1.9e-132;
Matches 579; Conservative 0; Mismatches 4; Indels 21; Gaps 1;
Qy 51 AGTGTGATGTTTCACTGGGATCTTCCCTTCGACCCCTTTGGCCACAGTGTGTGACCGCTGG 110
Db 1 AGTGTGATGTTTCACTGGGATCTTCCCTTCGACCCCTTTGGCCACAGTGTGTGACCGCTGG 60
Qy 111 GGACTGTGAGAGTGTGAGGGGACGTTCCAGCGCTGTGACTCTTCTCTCTCTACTAGAG 170
Db 61 GGACTGTGAGAGTGTGAGGGGACGTTCCAGCGCTGTGACTCTTCTCTCTCTACTAGAG 120
Qy 171 ACGCAGCCTATAGCTCCGCGCAGGCGAGTCTCCAGGAACTGAAATAGTGAATATAGATT 230
Db 121 ACGCAGCCTATAGCTCCGCGCAGGCGAGTCTCCAGGAACTGAAATAGTGAATATAGATT 180
Qy 231 GGCAGGAGAGATCAACATATAGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAG 290
Db 181 GGCAGGAGAGATCAACATATAGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAGGCTAG 240
Qy 291 TTGGGGCTATCTTACTGGCTCCCTTTGTCCAGGAACTGATGATGAAGAGCCTAAAG 350
Db 241 TTGGGGCTATCTTACTGGCTCCCTTTGTCCAGGAACTGATGATGAAGAGCCTAAAG 279
Qy 351 AAGGAGAAACCCACCTAAAGTCGGAATCTACACCTGATCAGAGAGAGAGATGATC 410
Db 280 AAGGAGAAACCCACCTAAAGTCGGAATCTACACCTGATCAGAGAGAGAGATGATC 339
Qy 411 AGGCTGAGCTGAGATTCAGTGCCTGACCTGGAAGCGGATCTCCAGGAGCTATGTGAGA 470
Db 340 AGGCTGAGCTGAGATTCAGTGCCTGACCTGGAAGCGGATCTCCAGGAGCTATGTGAGA 399
Qy 471 CAAGAGCTGGGAGTGGATGTGAAGTGGTACTGTGATGTCAGGGGAGAGATCTTACCAAAAG 530
Db 400 CAAGAGCTGGGAGTGGATGTGAAGTGGTACTGTGATGTCAGGGGAGAGATCTTACCAAAAG 459
Qy 531 CAGAGCCTTTAAATGCCAGACAGTGAAGGGAATCACAGGTTTAAAGGAAGATAA 590
Db 460 CAGAGCCTTTAAATGCCAGACAGTGAAGGGAATCACAGGTTTAAAGGAAGATAA 519
Qy 591 GCTGAAACACACAACTGTTTTTATATTAGATATTTTACCTTTAAAGAGTCTTAATAAT 650
Db 520 GCTGAAACACACAACTGTTTTTATATTAGATATTTTACCTTTAAAGAGTCTTAATAAT 579
Qy 651 TTTT 654
Db 580 TTTT 583

RESULT 5
BE7323213
LOCUS BG623213
DEFINITION BG623213.1 GI:13674584
ACCESSION BG623213
VERSION BG623213.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 652)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
```


Plate: L1CM1635	row: a	column: 08
High quality sequence stop: 627.		
Location/Qualifiers		
1. .652		
/organism="Homo sapiens"		
/mol_type="mRNA"		
/db_xref="taxon:9606"		
/clone="IMAGE:4769551"		
/lab_host="DH10B (TI phage-resistant)"		
/clone_lib="NIH_MGC_79"		
<p>Notes: Organ: placenta; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccgctggcc); Site 2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGCGCATTTATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGAGGCGCCGACATG-TT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."</p>		
BASE COUNT	192 a	144 c 165 g 151 t
ORIGIN		
Query Match	80.7%;	Score 540; DB 10; Length 652;
Best Local Similarity	95.4%;	Pred. No. 5.4e-131;
Matches 607;	Conservative 0;	Mismatches 5; Indels 24; Gaps 4;
Qy	22	AGTGTGCATGTTCTCTTGACACCTGCTTCAGTGTGCATGTTCACTTGGGCATCTTCCCTTC 81
Db	2	AGTGTGCATGTTCTCTTGACACCTGCTTCAGTGTGCATGTTCACTTGGGCATCTTCCCTTC 61
Qy	82	GACCCCTTGGCCACGTTGTCACCGCTGGGAGCTGTGAGAGTGTGAGGGGCAGCTTCCA 141
Db	62	GACCCCTTGGCCACGTTGTCACCGCTGGGAGCTGTGAGAGTGTGAGGGGCAGCTTCCA 121
Qy	142	CCCGTCTGGACTCTTTCTCTCTACTGAGACGCGACCTTATAGGTCCGAGCGCAGTCTCTC 201
Db	122	CCCGTCTGGACTCTTTCTCTCTACTGAGACGCGACCTTATAGGTCCGAGCGCAGTCTCTC 181
Qy	202	CCAGGACTGAAATAGTCAAAATATAGATTGGCGAGGAGAGATCAACATATAGCGCTAGGCC 261
Db	182	CCAGGACTGAAATAGTCAAAATATAGATTGGCGAGGAGAGATCAACATATAGCGCTAGGCC 241
Qy	262	AAGAAGAAAGTTTACAGCCTCTGAGCTGATTGGGCTATGCTTACTGTGCTCCCTTTGTC 321
Db	242	AAGAAGAAAGTTTACAGCCTCTGAGCTGATTGGGCTATGCTT----- 284
Qy	322	CCAGGAACCCACTGATGAAGAGCCCTAAAGAGAGAGAAACACCCACTAAAGTCGGAATCC 381
Db	285	-----GACCCACTGATGAAGAGCCCTAAAGAGAGAGAAACACCCACTAAAGTCGGAATCC 340
Qy	382	TACACCTGATCAGAAGAGAGAA -GATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACC 440
Db	341	TACACCTGATCAGAAGAGAGAAACGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACC 400
Qy	441	TGGAAGCCGATCTCCAGGAGCTA-TGTTCAGACAAAGACTGGGCGATGGATGTCAAGTGGT 499
Db	401	TGGAAGCCGATCTCCAGGAGCTATTGTCAGACAAAGACTGGGCGATGGATGTGAAGTGGT 460
Qy	500	ACTGATGTCAAGGGGAAGATTCTACAAAGCAGAGCACTTTAAATGCCAGAACGAGGT 559
Db	461	ACTGATGTCAAGGGGAAGATTCTACCAAAAGCAGAGCACTTTAAATGCCAGAACGAGGT 520
Qy	560	GAAAGGGAATCACAGTTTTAAAGGAAGATAAGCTGAAACAAACACAACTGTTTTTATA-T 618
Db	521	GAAAGGGAATCACAGTTTTAAAGGAAGATAAGCTGAAACAAACACAACTGTTTTTATACT 580
Qy	619	TAGATATTTTATCTTTAAAGAGTCTTAAATTTT 654
Db	581	TAGATATTTTATCTTTAAATATCTTAAATTTT 616

RESULT 6	AM972530/c	592 bp	mrna	linear	EST 01-JUN-2000
LOCUS	EST384621	WAGE resequences	MAGL Homo sapiens	cdna	mrna sequence.
DEFINITION	AM972530				
ACCESSION	AM972530.1	GI:8162376			
VERSION	EST.				
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1. (bases 1 to 592) Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and Quackenbush, J.				
TITLE	Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray				
JOURNAL	Unpublished				
COMMENT	Contact: John Quackenbush The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 3528 Fax: 301 838 0208 Email: johnq@tigr.org Plate: 307				
FEATURES	Seq primer: Forward.				
source	Location/Qualifiers				
	1..592				
	/organism="Homo sapiens"				
	/mol_type="mrna"				
	/db_xref="taxon:9606"				
	/clone_lib="WAGE resequences, MAGL"				
	/note="vector: pBluescriptSKm"				
BASE COUNT	138 a 152 c 122 g 179 t				1 others
ORIGIN					
Query Match	80.5%;	Score 538.4;	DB 10;	Length 592;	
Best Local Similarity	95.3%;	Pred. No. 1.4e-130;			
Matches	573;	Conservative 0;	Mismatches 7;	Indels 21;	Gaps 1
QY	54	GTGCATGTTTCACCTGGGCATCTTCCTTCGACCCCTTTGCCACCGTGTGACCGTGGGGA	113		
Db	591	GTGCATGTTTCATTGGGCATCTTCACCTTCGACCCCTTTGCCACCGTGTGACCGTGGGGA	532		
QY	114	GCTGTGAGAGTGTGAGGGGACGTTCCAGCGCTGTGACTCTTCTCTCCTACTGAGACG	173		
Db	531	NCTGTGAGAGTGTGAGGGGACGTTCCAGCGCTGTGACTCTTCTCTCCTACTGAGACG	472		
QY	174	CAGCCTATAGTTCGCGAGGCAGTCTCTCCAGGAACCTGAAATAGTGAATATGAGTTGGC	233		
Db	471	CAGCCTATAGTTCGCGAGGCAGTCTCTCCAGGAACCTGAAATAGTGAATATGAGTTGGC	412		
QY	234	GAGGAAGATCAACATATAGGCGTAGGCCAAGAAGTTTACAGCCTCTCTGAGCTGATTG	293		
Db	411	GAGGAAGATCAACATATAGGCGTAGGCCAAGAAGTTTACAGCCTCTCTGAGCTGATTG	352		
QY	294	GGGCTATGCTTACTGGCTCCCTTTGTGCCAGGAACCCACTGATGAAGAGCCTAAAGAG	353		
Db	351	GGGCTATGCTT-----GAACCCACTGATGAAGAGCCTAAAGAG	313		
QY	354	AGAAACCCACCTAAAGTCGGATCCTACACCTGATCAGAAGAGAGAAGATGATCAGG	413		
Db	312	AGAAACCCACCTAAAGTCGGATCCTACACCTGATCAGAAGAGAGAAGATGATCAGG	253		
QY	414	GTGCAGCTGAGATTCAGAGTGTGACCTGGAGCCGATCTCCAGGAGCTATGTCAGACAA	473		
Db	252	GTGCAGCTGAGATTCAGAGTGTGACCTGGAGCCGATCTCCAGGAGCTATGTCAGACAA	193		
QY	474	AGACTGGGGATGATGTCAGAGTGTGATGTCAGAGGGGAAGATTCTACCAAAAGCAG	533		
Db	192	AGACTGGGGATGATGTCAGAGTGTGATGTCAGAGGGGAAGATTCTACCAAAAGCAG	133		
QY	534	AGCACTTTAAATGCCAAGACGAGTGAAGGAAATCAACAGTTTAAAGGAAGATAAGCT	593		

of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 240 a 155 c 218 g 169 t

Query Match 79.4%; Score 531; DB 10; Length 782;
Best Local Similarity 94.6%; Pred. No. 1.3e-128;
Matches 591; Conservative 0; Mismatches 10; Indels 24; Gaps 3;

QY 31 GTTCTTGGACACCTGCTCAGTGTGATCTTCACTGGGCATCTTCCCTCGACCCCTTT 90
Db 2 GTTCTTGGACACCTGCTCAGTGTGATCTTCACTGGGCATCTTCCCTCGACCCCTTT 61
QY 91 GCCACATGCTGACCGCTGGGGAGCTGTGAGAGTGTGAGGGGACAGTTCAGCGCTCG 150
Db 62 GCCACATGCTGACCGCTGGGGAGCTGTGAGAGTGTGAGGGGACAGTTCAGCGCTCG 121
QY 151 ACTCTTTCTCTCTACTGACCGCAGCTATAGTTCGGCAGGCGAGTTCCTCCAGGAAC 210
Db 122 ACTCTTTCTCTCTACTGACCGCAGCTATAGTTCGGCAGGCGAGTTCCTCCAGGAAC 181
QY 211 GAAATAGTGAATATGAGTTGGGAGGAGGATCAACATATAGGCTAGGCCAAGAGAG 270
Db 182 GAAATAGTGAATATGAGTTGGGAGGAGGATCAACATATAGGCTAGGCCAAGAGAG 241
QY 271 TTACAGCCTCTGAGCTGATTTGGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAAC 330
Db 242 TTACAGCCTCTGAGCTGATTTGGGGCTATGCTT-----GAAC 280
QY 331 CACTGATGAAGAGCTTAAAGAGAGAACCAACCCACTTAAAGTCGGAATCTACACTGA 390
Db 281 CACTGATGAAGAGCTTAAAGAGAGAACCAACCCACTTAAAGTCGGAATCTACACTGA 340
QY 391 TCAGAGAGAGAGATGATCAGGGTGAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGA 450
Db 341 TCAGAGAGAGAGATGATCAGGGTGAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGA 400
QY 451 TCTCCAGAGCTATGTCAGACAAAGACTGGGGATGATGGAAGTGTACTGATGCAA 510
Db 401 TCTCCAGAGCTATGTCAGACAAAGACTGGGGATGATGGAAGTGTACTGATGCAA 460
QY 511 GGGAGAGTCTACCAAAAGCAGAGCACTTTAAATGCCAGAGCAGGTGAAGGGAATC 570
Db 461 GGGAGAGTCTACCAAAAGCAGAGCACTTTAAATGCCAGAGCAGGTGAAGGGAATC 520
QY 571 ACAGTTTAAAGGAGATAGCTGAAACACACAACTGTTTATATTAGATATTAC 630
Db 521 ACAGTTTAAAGGAGATAGCTGAAACACAACTGTTTATATTAGATATTAC 578
QY 631 TTTAAGA-GTCTTAAATAATTTT 654
Db 579 TTTAATAATCTTAAATAATTTT 603

RESULT 9
BE733225
LOCUS 601568462P1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3843231 5',
DEFINITION mRNA sequence.
ACCESSION BE733225
VERSION BE733225.1 GI:10147217
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 588)
NIH-MGC <http://mgi.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
AUTHORS Contact: Robert Strausberg, Ph.D.
TITLE Email: cgabs-r@mail.nih.gov
COMMENT Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LCM537 row: 1 column: 16
High quality sequence stop: 588.

FEATURES
Location/Qualifiers
1..588
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone IMAGE:3843231
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_21"
/note="Organ: placenta; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >50bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 178 a 124 c 152 g 134 t

ORIGIN
Query Match 75.7%; Score 506.6; DB 10; Length 588;
Best Local Similarity 95.3%; Pred. No. 3.2e-122;
Matches 573; Conservative 0; Mismatches 4; Indels 24; Gaps 4;

QY 54 GTGCATGTTCACTGGGCATCTTCCCTTCGACCCCTTTGCCACCGTGGACCGCTGGGGA 113
Db 1 GTGCATGTTCACTGGGCATCTTCCCTTCGACCCC-TTGCCACCGTGGACCGCTGGGGA 59
QY 114 GCTGTGAGAGTGTGAGGGGACGTTCCAGCCGTCTGGACTCTTTCTCTCTACTGAGACG 173
Db 60 GCTGTGAGAGTGTGAGGGGACGTTCCAGCCGTCTGGACTCTTTCTCTCTACTGAGACG 119
QY 174 CAGCCTATAGTTCGCGAGCGCAGTCCCTCCAGGAACTGAATATGAAATGAGTTGGC 233
Db 120 CAGCCTATAGTTCGCGAGCGCAGTCCCTCCAGGAACTGAATATGAAATGAG-TGGC 178
QY 234 GAGGAAGATCAACATATAGGCTAGGCCAAGAAAGTTTACAGCCTCCTGAGCTGATTG 293
Db 179 GAGGAAGATCAACATATAGGCTAGGCCAAGAAAGTTTACAGCCTCCTGAGCTGA-TG 237
QY 294 GGGCTATGCTTACTGGCTCCCTTTGTCAGGAAACCCTGATGAAGAGCTTAAGNAG 353
Db 238 GGGCTATGCTT-----GAACCCACTGATGAAGAGCTTAAGNAG 276
QY 354 AGAAACCCACCCACTAAAGTGGGAATCTACACCTGATCAGAGAGAGAGATGATCAGG 413
Db 277 AGAAACCCACCCACTAAAGTGGGAATCTACACCTGATCAGAGAGAGAGATGATCAGG 336
QY 414 GTGCAGCTGAGATTCAAGTGCCTGACCTGGAGCGGATCTCCAGAGCTATGTACAGAA 473
Db 337 GTGCAGCTGAGATTCAAGTGCCTGACCTGGAGCGGATCTCCAGAGCTATGTACAGAA 396
QY 474 AGACTGGGGATGATGTAAGTGGTACTGATGTCAAGGGGAAAGATTCTACCAAAAGCAG 533
Db 397 AGACTGGGGATGATGTAAGTGGTACTGATGTCAAGGGGAAAGATTCTACCAAAAGCAG 456
QY 534 AGCACTTTAAATGCCAGAGCAGGTGAAGGGAATCAACAGTTTAAAGGAAGATAAGCT 593
Db 457 AGCACTTTAAATGCCAGAGCAGGTGAAGGGAATCAACAGTTTAAAGGAAGATAAGCT 516
QY 594 GAAACACACAACTGTTTTTATATAGATATTTTACTTTTAAAGAGTCTTAAATAATTTT 653
Db 517 GAAACACACAACTGTTTTTATATAGATATTTTACTTTTAAATAATCTTAAATAATTTT 576
654 T 654

ORIGIN

Query Match	69.4%	Score	464.4	DB	13	Length	498		
Best Local Similarity	95.8%	Pred.	No. 3.8e-111						
Matches	496	Conservative	0	Mismatches	1	Indels	21	Gaps	1

QY	31	GTTCCTTTGGACACCTGCCTCAGTGTGATGTTCACTGGGCACTCTCCCTTCGACGCCCTTT	90
Db	2	GTTCCTTTGGACACCTGCCTCAGTGTGCAATGTTCACTGGGCATCTTCCTTCGACCCCTTT	61
QY	91	GCCACAGTGTGACCGCTGGGGAGCTGTGAGAGTGTGAGGGGCAAGTTTCAGCCGCTCTGG	150
Db	62	GCCACAGTGTGACCGCTGGGGAGCTGTGAGAGTGTGAGGGGCAAGTTTCAGCCGCTCTGG	121
QY	151	ACTCTTTCTCTCTCTACTGAGACGCGACCTATAGTGTCCGCGAGGCGAGTCTCCACGGAAC	210
Db	122	ACTCTTTCTCTCTCTACTGAGACGCGACCTATAGTGTCCGCGAGGCGAGTCTCCACGGAAC	181
QY	211	GAATAGTGAATATGATGTTGGCGAGGAGATCAACATATATAGGCGCTAGGCCCAAGAAGAG	270
Db	182	GAATAGTGAATATGATGTTGGCGAGGAGATCAACATATATAGGCGCTAGGCCCAAGAAGAG	241
QY	271	TTTACAGCCTCTCAGAGCTGATTGGGGGCTATGCTTACTGGCTCCCTTTGTCCACGGAACC	330
Db	242	TTTACAGCCTCTCAGAGCTGATTGGGGGCTATGCTT-----GAAACC	280
QY	331	CACGTATGAGAGCGCTTAAGAGAGGAACCAACCCACTAAAAGTCGGAATCCTACACCTGA	390
Db	281	CACGTATGAGAGCGCTTAAGAGAGGAACCAACCCACTAAAAGTCGGAATCCTACACCTGA	340
QY	391	TCAGAAGAGAGAAGATGATCAGGGGTGCAGGCTGAGATTCAAGTGCCTGACCTGGGAAGCCGA	450
Db	341	TCAGAAGAGAGAAGATGATCAGGGGTGCAGGCTGAGATTCAAGTGCCTGACCTGGGAAGCCGA	400
QY	451	TCTCCAGAGCTATGTGAGACAAAGACTGGGGATGCGATGTGAAGGTGGTACTGATGTCAA	510
Db	401	TCTCCAGAGCTATGTGAGACAAAGACTGGGGATGCGATGTGAAGGTGGTACTGATGTCAA	460
QY	511	GGGGAAGATTCTACCAAAAGCAGAGCACTTTTAAATGC	548
Db	461	GGGGAAGATTCTACCAAAAGCAGAGCACTTTTAAATGC	498

```

RESULT 12
LOCUS      T87105/c
DEFINITION T87105 Homo sapiens CDNA clone
            533 bp mRNA linear EST 17-MAR-1995
            v488e11.1 Soares fetal liver spleen lNFLS Homo sapiens
            IMAGE115340 3', mRNA sequence.
ACCESSION T87105
VERSION   T87105.1 GI:715457
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 533)
AUTHORS   Hallier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            M.J., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston
            R., Williamson,A., Woldmann,P. and Wilson,R.
TITLE     The Washu-Merck EST Project
JOURNAL   Unpublished
COMMENT   Contact: Wilson RK
            Washington University School of Medicine
            444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 449
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -21m13

```

High quality sequence stop: 449.

FEATURES
SOURCE

BASE COUNT	121 a	128 c	105 g	177 t	2 others	
ORIGIN						
Query Match	66.5%;	Score 445.2;	DB 14;	Length 533;		
Best Local Similarity	93.8%;	Pred. No. 4.4e-106;				
Matches 503; Conservative	0;	Mismatches 10;	Indels 23;	Gaps 3;		
QY	120	AGAGTGTGAGGGGCACGTTCCAGCCGTCGTGGACTCTTTCTCTCT-CTTACTGAGACGCGAGCC	178			
DB	533	AAATTTTNNAGGGGCACGTTCTNAGCCGTCGTGGACTCTTTCTCTCCTACTGAGACGCGAGCC	474			
QY	179	TATAGTCCGCGAGGCGCAGTCCCTCCAGGAACCTCAAAATAGTGAATATGAGTTGGCGAGGA	238			
DB	473	TATAGTCCGCGAGGCGCAGTCCCTCCGGAACCTGAAATAGTGAATATGAGTTGGCGAGGA	414			
QY	239	AGATCAACATATATAGGCGCTAGGCCAAGAAGTTTACAGCCTCCTGAGCTTGATTTGGGGCT	298			
DB	413	AGATCAACATATATA-GCCTAGGCCAAGAAGTTTACAGCCTCCTGAGCTTGATTTGGGGCT	355			
QY	299	ATGCTTTACTGGCTCCCTCTTTGTCCAGGAACCCACCTGATGAAGAGCGCTTAAGAAGAGAGAAA	358			
DB	354	ATGCTT-----GAACCCACTGATGAAGAGCGCTTAAGAAGAGAGAAA	316			
QY	359	CAACCCACTAAAGTCGGAATCTTACACCTGATCAGAGAAGAGAAGATGATCAGGGTGCA	418			
DB	315	CAACCCACTAAAGTCGGAATCTTACACCTGATCAGAGAAGAGAAGATGATCAGGGTGCA	256			
QY	419	GCTGAGATTCAAGTGCGCTGACCTGGAAGCGGATCTCCAGGAGCTATGTCCAGACAAAGACT	478			
DB	255	GCTGAGATTCAAGTGCGCTGACCTGGAAGCGGATCTCCAGGAGCTATGTCCAGACAAAGACT	196			
QY	479	GGGGATGGAATGGAAGGTGGTATGATGTCAAGGGGAAGATTCTACAAAAGCAGAGCAC	538			
DB	195	GGGGATGGAATGGAAGGTGGTATGATGTCAAGGGGAAGATTCTACAAAAGCAGAGCAC	136			
QY	539	TTTTAAATGCCAAGCAGGTGAAGGGAATCACAGSTTTAAAGGAAGATAAGCTGAAC	598			
DB	135	TTTTAAATGCCAAGCAGGTGAAGGGAATCACAGSTTTAAAGGAAGATAAGCTGAAC	76			
QY	599	AACACAAACGTTTTTATATATAGATATTTTACATTTTAAAGAGTCTTAATAAATTTT	654			
DB	75	AACACAAACGTTTTTATATATAGATATTTTACTTTTAAATATCTTAATAAGTTTT	20			

RESULT 13	AA9931199/c	496 bp	mRNA	linear	EST 27-AUG-1998
LOCUS	AA9931199/c				
DEFINITION	ct92c08 s1 Soares total fetal brain cDNA clone				
	IMAGE:1524334.3		similar to SW:GGE8_HUMAN Q13066		GAGE-2 PROTEIN. [1
	1 : mRNA sequence.				
ACCESSION	AA9931199				
VERSION	AA9931199.1				
	GI:3179744				

```

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE 1 (bases 1 to 496)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 634 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amerham
High quality sequence stop: 474.
FEATURES
Location/Qualifiers
1..496
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1624334"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/clone_lib="Soares total_fetus_Nb2HF8_9w"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand
was prepared from mRNA obtained from pooled 8-9 week
(totally) fetus material with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGGAGCGCGCGCTAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 112 a 119 c 96 g 169 t
ORIGIN
Query Match 66.2%; Score 442.6; DB 9; Length 496;
Best Local Similarity 95.0%; Pred. No. 2.1e-105;
Matches 476; Conservative 0; Mismatches 4; Indels 21; Gaps 1;
QY 154 CTTCTCTCTCTACTGAGCGGAGCCTATAGTGCAGCGCCAGTCTCTCCAGGAACTGAA 213
DB 496 CTTCTCTCTCTACTGAGCGGAGCCTATAGTGCAGCGCCAGTCTCTCCAGGAACTGAA 437
QY 214 ATAGTGAATATAGTTCGCGAGGAGATCAACATATAGGCTAGGCCCAAGAAAGTTT 273
DB 436 ATAGTGAATATAGTTCGCGAGGAGATCAACATATAGGCTAGGCCCAAGAAAGTTT 377
QY 274 ACAGCTCTGAGCTGATTCGGGCTATGCTTACTGCTCCCTTTGTCCTCCAGGACCCAC 333
DB 376 ACAGCTCTGAGCTGATTCGGGCTATGCTTACTGCTCCCTTTGTCCTCCAGGACCCAC 338
QY 334 TGATGAAGAGCCTTAAAGAGAGAAACCCACCTAAAGTCGGAATCTTACACCTGATCA 393
DB 337 TGATGAAGAGCCTTAAAGAGAGAAACCCACCTAAAGTCGGAATCTTACACCTGATCA 278
QY 394 GAAGAGAGAGATGATCAGGTTCAGCTGAGATTCAGGTTCAGCTGAGGCGGATCT 453
DB 277 GAAGAGAGAGATGATCAGGTTCAGCTGAGATTCAGGTTCAGCTGAGGCGGATCT 218
QY 454 CCAGGAGCTATGTTCAGCAAAAGACTGGGATGATGTGAAGGTGGTACTGATGTCAAGGG 513
DB 217 CCAGGAGCTATGTTCAGCAAAAGACTGGGATGATGTGAAGGTGGTACTGATGTCAAGGG 158
QY 514 GAAGATTCTACAAAAGCAGAGACACTTAAATGCCAGAGCAGGTGAAGGGAATCACA 573
DB 157 GAAGATTCTACAAAAGCAGAGACACTTAAATGCCAGAGCAGGTGAAGGGAATCACA 98
QY 574 GGTTTAAAGAGATAGCTGAAACCAACAACTGTTTTTATATAGATATTTTACTTTT 633

```

```

Db 97 GGTTTAAAGAGATAGCTGAAACCAACAACTGTTTTTATATATAGATATTTTACTTT 38
QY 634 AAAGAGCTCTTAATAAATTTT 654
Db 37 AAAATATCTTTAATAAAGTTT 17
RESULT 14
LOCUS A1188535/c
DEFINITION 496 bp mRNA linear EST 28-OCT-1998
cd09e11.x1 Soares placenta_8to9weeks_2NHP8to9w Homo sapiens cDNA
clone IMAGE:1723244_3' similar to SW:GGE2_HUMAN Q13066 GAGE-2
PROTEIN. [1]; mRNA sequence.
ACCESSION A1188535
VERSION A1188535.1 GI:3739744
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 601 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 473.
FEATURES
Location/Qualifiers
1..496
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1723244"
/dev_stage="two placenta: one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares placenta_8to9weeks_2NHP8to9w"
/note="Organ: placenta; Vector: p773D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGGAGCGCGCGATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p773 vector
(Pharmacia). Library constructed by Bento Soares and
M. Fatima Bonaldo."
BASE COUNT 109 a 119 c 95 g 173 t
ORIGIN
Query Match 65.6%; Score 438.6; DB 9; Length 496;
Best Local Similarity 95.0%; Pred. No. 2.3e-104;
Matches 472; Conservative 0; Mismatches 4; Indels 21; Gaps 1;
QY 158 CTCTCTCTACTGAGACGCGCCTATAGTTCGCGAGCGCAGTCTCTCCAGGAACTGAAATAG 217
DB 496 CTCTCTCTACTGAGACGCGCCTATAGTTCGCGAGCGCAGTCTCTCCAGGAACTGAAATAG 437
QY 218 TGAATATAGTTCGCGAGGAAATCAACATATAGGCTAGGCCCAAGAGAGAGTTTACAG 277
DB 436 TGAATATAGTTCGCGAGGAAATCAACATATAGGCTAGGCCCAAGAGAGAGTTTACAG 377
QY 278 CCTCTGAGCTGATTCGGGCTATGCTTACTGCTCCCTTTGTCCTCCAGGACCCACTGAT 337
DB 376 CCTCTGAGCTGATTCGGGCTATGCTTACTGCTCCCTTTGTCCTCCAGGACCCACTGAT 338
QY 338 GAAGAGCTTAAAGAGAGAAACCCACCTAAAGTCGGAATCTTACACCTGATCTGAGAG 397
DB 337 GAAGAGCTTAAAGAGAGAAACCCACCTAAAGTCGGAATCTTACACCTGATCTGAGAG 278

```


/db xref="taxon:9606"
 /clone="IMAGE:1715402"
 /dev stage="two placentae: one from 8 weeks and another
 from 9 weeks post conception"
 /lab host="PH103 (ampicillin resistant)"
 /clone_lib="Soares_placenta_25weeks_2NHP8to9W"
 /note="Organ: placenta; Vector: pT73D (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'
 TGTTACCATCTGAAGTGGGAGCGCGCGGATTTTTTTTTTTT 3']
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT73 vector
 (Pharmacia). Library constructed by Bento Soares and
 M. Fatima Bonaldo."

BASE COUNT 105 a 119 c 90 g 169 t

Query Match 63.8%; Score 426.6; DB 9; Length 483;
 Best Local Similarity 94.6%; Pred. No. 3.3e-101;
 Matches 460; Conservative 0; Mismatches 4; Indels 21; Gaps 1;

QY 170 GAGCGAGCTATAGTCCGAGCCAGTCTCCAGGAACTGAATATAGTGAATATAGT 229
 Db 483 GAGCGAGCTATAGTCCGAGCCAGTCTCCAGGAACTGAATATAGTGAATATAGT 424
 QY 230 TGCGGAGGAGATCAACATATAGGCTAGGCCCAAGAGAGTTTACAGCTCTCTGAGCTG 289
 Db 423 TGCGGAGGAGATCAACATATAGGCTAGGCCCAAGAGAGTTTACAGCTCTCTGAGCTG 364
 QY 290 ATTGGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACTGATGAAGGCTTAAA 349
 Db 363 ATTGGGGCTATGCTT-----GAACCCACTGATGAAGGCTTAAA 325
 QY 350 GAAGAGAAACCCCACTAAAAGTGGAACTCTACACCTGATCAGAGAGAGAGATGAT 409
 Db 324 GAAGAGAAACCCCACTAAAAGTGGAACTCTACACCTGATCAGAGAGAGAGATGAT 265
 QY 410 CAGGGTGACGCTGAGATTTCAAGTGGCTGACCTGGAGCGGATCTCCAGGAGCTATGTCAG 469
 Db 264 CAGGGTGACGCTGAGATTTCAAGTGGCTGACCTGGAGCGGATCTCCAGGAGCTATGTCAG 205
 QY 470 ACAAGACTGGGATGGATGTGAAGTGGTACTGTGTCAAGGGGAAGATTTCTACCAAAA 529
 Db 204 ACAAGACTGGGATGGATGTGAAGTGGTACTGTGTCAAGGGGAAGATTTCTACCAAAA 145
 QY 530 GCAGGCACTTTAAATGCGAGAGCAGGTGAGGGAATCAGAGTTTAAAGGAAGATA 589
 Db 144 GCAGGCACTTTAAATGCGAGAGCAGGTGAGGGAATCAGAGTTTAAAGGAAGATA 85
 QY 590 AGCTGAACCAACCAAACTGTTTTATATTAGATATTTTACTTTTAAAGAGCTTTTAAATAAA 649
 Db 84 AGCTGAACCAACCAAACTGTTTTATATTAGATATTTTACTTTTAAAGAGCTTTTAAATAAA 25
 QY 650 TTTT 654
 Db 24 GTTT 20

RESULT 17
 A1128458/c
 LOCUS
 DEFINITION
 IMAGE:1711966 3', similar to SW:GG2_HUMAN Q13066 GAGE-2 PROTEIN. [1
]; mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Human sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 468)

AUTHORS
 TITLE
 JOURNAL
 COMMENT

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 596 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 444.

FEATURES
 source

1..468
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1711966"
 /sex="female"
 /lab_host="DH10B"
 /dev stage="adult"
 /clone_lib="Soares_pregnant_uterus_NbHPU"
 /note="Organ: uterus; Vector: pT73-Pac; Site 1: Not I;
 Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer [5'
 AACTGGAAGATTCGCGCGCCCTTTTTTTTTTTT 3']
 double-stranded cDNA was ligated to Eco RI adapters
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization. Library
 constructed by M. Fatima Bonaldo."

BASE COUNT 102 a 111 c 83 g 172 t

Query Match 60.5%; Score 404.6; DB 9; Length 468;
 Best Local Similarity 94.6%; Pred. No. 2e-95;
 Matches 438; Conservative 0; Mismatches 4; Indels 21; Gaps 1;

QY 192 GCAGTCTCTCCAGGAACTGAATAGTGAATATAGTTCGCGAGGAGATCAACATATA 251
 Db 468 GCAGTCTCTCCAGGAACTGAATAGTGAATATAGTTCGCGAGGAGATCAACATATA 409
 QY 252 GGCCTAGGCAAGAGAGTTTACAGGCTCTCTGAGCTGATTTGGGGCTATGCTTACTGGCT 311
 Db 408 GGCCTAGGCAAGAGAGTTTACAGGCTCTCTGAGCTGATTTGGGGCTATGCTT----- 356
 QY 312 CCCCTTTGTCGCCAGGACCCACTGATGAAGCCCTTAAGAGAGCAACCCCACTAAAA 371
 Db 355 -----GAACCCACTGATGAAGAGCCCTTAAGAGAGCAACCCCACTAAAA 310
 QY 372 GTCGGAATCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAAGCTGAGATTCAAG 431
 Db 309 GTCGGAATCCTACACCTGATCAGAGAGAGAGATGATCAGGGTGCAAGCTGAGATTCAAG 250
 QY 432 TGCCCTGACCTGGAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGGATGGATGG 491
 Db 249 TGCCCTGACCTGGAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGGATGGATGG 190
 QY 492 AAGGTGGTACTGATGTCAAGGGGAAGATTCTACAAAAGCAGAGCACTTTTAAATGCCAG 551
 Db 189 AAGGTGGTACTGATGTCAAGGGGAAGATTCTACAAAAGCAGAGCACTTTTAAATGCCAG 130
 QY 552 AAGCAGGTGAAGGGGAATCAGAGCTTTAAAGGAGATTAAGCTGAAACACCAAACTGTT 611
 Db 129 AAGCAGGTGAAGGGGAATCAGAGCTTTAAAGGAGATTAAGCTGAAACACCAAACTGTT 70
 QY 612 TTTATATTAGATATTTTACTTTTAAAGAGCTCTTAATAAATTTT 654
 Db 69 TTTATATTAGATATTTTACTTTTAAATATCTTAATAAAGTTT 27

RESULT 18
 AA039331/c
 LOCUS

AA039331 468 bp mRNA linear EST 10-MAY-1997

```

zk39e06.s1 Soares_pregnant uterus_NbHPU Homo sapiens cDNA clone
IMAGE:485218 3', mRNA sequence.
ACCESSION      AA039331
VERSION        AA039331.1 GI:1515609
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
REFERENCE      1 (bases 1 to 468)
AUTHORS        Hillier, L., Lennon, G., Becker, M., Bonaldo, M.P., Chiapelli, B.,
               Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, N., Hawkins
               , B., Hultman, M., Kucaba, F., Lacy, M., Le, M., Le, N., Marid, E., Moore
               , B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,
               Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaaskis, E.,
               Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.
               Generation and analysis of 280,000 human expressed sequence tags
               Genome Res. 6 (9), 807-828 (1996)
TITLE          97044478
JOURNAL        MEDLINE
PUBMED         8889549
COMMENT        Contact: Wilson RK
               Washington University School of Medicine
               4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
               Tel: 314 286 1800
               Fax: 314 286 1810
               Email: estowatson.wustl.edu
               This clone is available royalty-free through LLNL ; contact the
               IMAGE Consortium (info@image.llnl.gov) for further information.
               Insert Length: 597 Std Error: 0.00
               Seq primer: -40M13 fwd. from Amersham
               High quality sequence stop: 398.
               Location/Qualifiers
                 1..468
                   /organism="Homo sapiens"
                   /mol_type="mRNA"
                   /db_xref="GDB:3758932"
                   /db_xref="taxon:9606"
                   /clone="IMAGE:485218"
                   /sex="female"
                   /dev stage="adult"
                   /lab_host="DH10B"
                   /clone_lib="Soares_pregnant_uterus_NbHPU"
                   /notes="Organ: uterus; Vector: p7T73-Pac; Site 1: Not I;
                   Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
                   oligo(dT) primer [5',
                   AACTGGAAGAATTTCGGCGCGCTTTTTTTTTTTTTTTT 3'],
                   double-stranded cDNA was ligated to Eco RI adaptors
                   (Pharmacia), digested with Not I and cloned into the Not I
                   and Eco RI sites of the modified p7T73 vector. Library
                   went through one round of normalization. Library
                   constructed by M. Fatima Bonaldo."
BASE COUNT    102 a 113 c 87 g 166 t
ORIGIN
Query Match   60.4%; Score 404; DB 9; Length 468;
Best Local Similarity 94.4%; Pred. No. 2.9e-95;
Matches 438; Conservative 0; Mismatches 5; Indels 21; Gaps 1;

QY 191 GGCCAGTCTCCAGGAACCTGAATAGTGAATATAGTTGGCGAGGAAGATCAACATAT 250
DB 462 GCCCAGTCTCCAGGAACCTGAATAGTGAATATAGTTGGCGAGGAAGATCAACATAT 403
QY 251 AGCCCTAGGCCAGAGAGATTTCAGCCCTCTGAGCTGATTGGGCTATGCTTACTGGC 310
DB 402 AGCCCTAGGCCAGAGAGATTTCAGCCCTCTGAGCTGATTGGGCTATGCTTACTGGC 349
QY 311 TCCCTCTTTGTCGCCAGAACCCACTGATGAAGAGCCCTAAAGAGAGAACCCACTTAA 370
DB 348 -----GNACCCACTGATGAAGAGCCCTAAAGAGAGAACCCACTTAA 304
QY 371 AGTCGGAATCCTACACTGATCAAGAGAGAGAGATGATCAGGGTSCAGCTCAGATTCAA 430
DB 303 AGTCGGAATCCTACACTGATCAAGAGAGAGAGATGATCAGGGTSCAGCTCAGATTCAA 244

```

```

QY 431 GTGCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGGATGT 490
DB 243 GTGCTGACCTGGAAGCCGATCTCCAGGAGCTATGTCAGACAAAGACTGGGATGGATGT 184
QY 491 GAAGTGCTGCTACTGATGTCAAGGGGAAGATTCTACAAAAGCAGAGCACTTTAAATGCCA 550
DB 183 GAAGTGCTGCTACTGATGTCAAGGGGAAGATTCTACAAAAGCAGAGCACTTTAAATGCCA 124
QY 551 GAACGAGTGAAGGGAATCACAGGTTTAAAGGAAGTAAGCTGAACACACAAACTGT 610
DB 123 GAACGAGTGAAGGGAATCACAGGTTTAAAGGAAGTAAGCTGAACACACAAACTGT 64
QY 611 TTTTATATTAGATATTTTACTTTTAAAGAGTCTTTAATAAATTTT 654
DB 63 TTTTATATTAGATATTTTACTTTTAAATAATCTTTAATAAATTTT 20

RESULT 19
A1189715/c
LOCUS      A1189715
DEFINITION Qd28e08.x1 Soares_placenta 8to9weeks 2NbHP8to9W Homo sapiens cDNA
            clone IMAGE:1725062 3' similar to SW:GGE2_HUMAN Q13066 GAGE-2
            PROTEIN. [1] ; mRNA sequence.
ACCESSION  A1189715
VERSION    A1189715.1 GI:3740924
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 453)
AUTHORS    NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL    Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: rcapbs-remail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 563 Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 434.
            Location/Qualifiers
              1..453
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:1725062"
                /dev stage="two placentae; one from 8 weeks and another
                from 9 weeks post conception"
                /lab host="DH10B (ampicillin resistant)"
                /clone_lib="Soares_placenta 8to9weeks 2NbHP8to9W"
                /notes="Organ: placenta; Vector: p7T73D (Pharmacia) with a
                modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
                strand cDNA was primed with a Not I - oligo(dT) primer [5',
                TGTACCAATCTGAAGTGGAGCGCGCGATTTTTTTTTTTTTTTT 3'],
                double-stranded cDNA was size selected, ligated to Eco RI
                adaptors (Pharmacia), digested with Not I and cloned into
                the Not I and Eco RI sites of a modified p7T73 vector
                (Pharmacia). Library constructed by Bento Soares and
                M. Fatima Bonaldo."
BASE COUNT 100 a 109 c 76 g 168 t
ORIGIN
Query Match   58.5%; Score 391.6; DB 9; Length 453;
Best Local Similarity 94.4%; Pred. No. 5.3e-92;
Matches 425; Conservative 0; Mismatches 4; Indels 21; Gaps 1;

QY 205 GGAACTGAATAGTGAATATGATTCGGCGAGGAAGATCAACATATAGGCTTAGGCCAAG 264
DB 453 GGAACTGAATAGTGAATATGATTCGGCGAGGAAGATCAACATATAGGCTTAGGCCAAG 394

```

BASE COUNT	127 a	128 c	107 g	169 t	1 others
ORIGIN					
Query Match	56.4%	Score 377.6	DB 9	Length 532	
Best Local Similarity	92.1%	Pred. NO. 2.7e-88			
Matches 417	Conservative 0	Mismatches 15	Indels 21	Gaps 1	
QY	215	TAGTGAATATGAGTTGGCGAGGAAGATCAACATATAGGCGCTAGGCCAAGAAGAGTTT	274		
DB	432	TAGTGAATATGAGTTGGCGAGGAAGATCAACATATAGGCGCTAGGCCAAGAAGAGTTT	373		
QY	275	CAGCCTCTGAGCTGATGGGCGTATGCTTACTGGCTCCCTTTGTCGCCAGGAAACCCACT	334		
DB	372	CAGCCTCTGAGCTGATGGGCGTATGCTTACTGGCTCCCTTTGTCGCCAGGAAACCCACT	334		
QY	335	GATGAAGAGCCCTAAAGAGAGAAACCAACCACTAAAAGTCGGAATCCTACACCTGATCAG	394		
DB	333	GATGAAGAGCCCTAAAGAGAGAAACCAACCACTAAAAGTCGGAATCCTACACCTGATCAG	274		
QY	395	AGAGAGAGAGATGATCAGGGTCGAGCTGAGATTCAGATGCTGCTGACCTGGAAGCCGATCTC	455		
DB	273	AAAGAGAGAGATGATCAGGGTCGAGCTGAGATTCAGATGCTGCTGACCTGGAAGCCGATCTC	214		
QY	455	CAGGAGCTATGTCAGACAAAGACTGGGGATGCGATGTGAAGGTGGTACTGATGTCGAAGGGG	514		
DB	213	CAGGAGCTATGTCAGACAAAGACTGGGGATGCGATGTGAAGGTGGTACTGATGTCGAAGGGG	154		
QY	515	AGATTCTTACCAAGAGCAGAGCACTTTTAAATGCCAGAGCAGGTGAAGGGAATCAACAG	574		
DB	153	AGATTCTTACCAAGAGCAGAGCACTTTTAAATGCCAGAGCAGGTGAAGGGAATCAACAG	94		
QY	575	GTTTAAAGGAAGATAAGCTGAAACACACAACTGTTTTTATATTAGATATTTTACTTTTA	634		
DB	93	GTTTAAAGGAAGATAAGCTGAAACACACAACTGTTTTTATATTAGATATTTTACTTTTA	34		
QY	635	AGAGGTCCTTAATAAATTTTTTGGCATGCTCGATC	667		
DB	33	CAATATCTTAATAAGTTTAAAGCTTTTCTGTC	1		
RESULT 21					
H53211			493 bp	mrna	linear
LOCUS	YG84C02.r1 Soares fetal liver spleen INFLS	Homo sapiens	cdna	clone	EST 20-SEP-1999
DEFINITION	IMAGE:202466 5', mRNA sequence.				
ACCESSION	H53211	GI:993358			
VERSION	H53211.1				
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 493) Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Woldmann, P., and Wilson, R.				
TITLE	The WashU-Merck EST Project				
JOURNAL	Unpublished				
COMMENT	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu Insert Size: 663 High quality sequence stops: 340 Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert length: 663 Std Error: 0.00 Seq primer: M3RPR1 High quality sequence stop: 340.				

Db 83 AAGATAAGCTGAACACACACAACTGTTTTATATATAGATATTTTACTTTAAATATCTT 24

QY 644 AATAAATTTT 654

Db 23 AATAAAGTTTT 13

RESULT 23

LOCUS W93043/c

DEFINITION ZD93f02.s1 Soares fetal heart_NbHL19W Homo sapiens cDNA clone IMAGE:357051 3', mRNA sequence.

ACCESSION W93043

VERSION W93043.1 GI:1422214

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1. (bases 1 to 512)

AUTHORS Hillier,L., Clark,N., Dubuque,T., Eliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.

TITLE The WashU-Werck EST Project

JOURNAL Unpublished.

COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 608 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 301.

FEATURES

Location/Qualifiers

1..512

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:1273595"

/db_xref="taxon:9606"

/clone="IMAGE:357051"

/sex="unknown"

/dev_stage="19 weeks"

/lab_host="DH10B (ampicillin resistant)"

/clone_lib="Soares fetal heart_NbHL19W"

/note="Organ: heart; Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCATCTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W."

BASE COUNT 119 a 121 c 108 g 162 t 2 others

ORIGIN

Query Match 55.6%; Score 372.2; DB 14; Length 512;

Best Local Similarity 90.8%; Pred. No. 6.9e-87;

Matches 471; Conservative 0; Mismatches 20; Indels 28; Gaps 6;

QY 143 CGCTCGGACTTTCTCT-CCTACTGAGAGC-CAGCCTATAGTCC--GCAGGCCAGTC 198

Db 511 CGGCTGGACTTTTCTCTCCCTACTGAGACGCCAGCCTATAGTCCCGCAGGCCAGTC 452

QY 199 CTCGCCAGGAACCTGAATAGTGAATATGATGGCGAGGAAGA--TCAACATATAGGCT 256

Db 451 CTCCTCCGGAACGAATAGTGAATATGATGGCGGAGGAAGATTCCACATATTGGCCT 392

QY 257 AGG-CCAGAGAAGCTTTACAGCCTCTGAGCTGATTTGGGCTATGCTTACTGGCTCCCC 315

Db 391 AGCCCCAAGAAGAAGTTTACAGCCTCTGAGCTGATTTGGGCTATGCTT-----343

QY 316 TTTGTCCCAAGAACCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCCTAAAGTCG 375

Db 342 -----GAACCCACTGATGAAGAGCCTAAAGAGAGAAACCCACCCTAAAGTCG 293

QY 376 GAATCCTACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGATGCC 435

Db 292 GAATCCTACCTGATCAGAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAGATGCC 233

QY 436 TGACCTGGAAGCCGATCTCCAGGAGCTATGTGAGACAAAGACTGGGGATGGATGTGAAGG 495

Db 232 TGACCTGGAAGCCGATCTCCAGGAGCTATGTGAGACAAAGACTGGGGATGGATGTGAAGG 173

QY 496 TGGTACTGATGTCAGAGGGAAGATTCTACCAAAAGCAGAGCACTTTAAATGCCGAGAGC 555

Db 172 TGGTACTGATGTCAGAGGGAAGATTCTACCAAAAGCAGAGCACTTTAAATGCCGAGAGC 113

QY 556 AGGTGAAGGGAATCACAGCTTTAAAGGAAGATAAGCTGAAACACACAAAAGCTGTTTTA 615

Db 112 AGGTGAAGGGAATCACAGCTTTAAAGGAAGATAAGCTGAAACACACAAAAGCTGTTTTA 53

QY 616 TATTAGATATTTTACTTTAAAGAGTCTTAATAAATTTT 654

Db 52 TATTAGATATTTTACTTTAAATATCTTTAAAGTTTT 14

RESULT 24

BE302172/c

LOCUS BE302172

DEFINITION BE302172.1 GI:9185920

445 bp mRNA linear EST 14-JUL-2000

similar to SW:GGE2_HUMAN Q13066 GAGE-2 PROTEIN. [1] ; mRNA sequence.

ACCESSION BE302172

VERSION BE302172.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1. (bases 1 to 445)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@ps-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml
Seq primer: -400P from Gibco
High quality sequence stop: 213.

FEATURES

Location/Qualifiers

1..445

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3048867"

/cell_line="MGC36"

/lab_host="DH10B"

/clone_lib="NIH MGC 10"

/note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 Kb. Library prepared by Life Technologies."

BASE COUNT 96 a 103 c 79 g 167 t

/organism="Homo sapiens"

/organism="Homo sapiens"

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 437)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Straube, Ph.D.
Email: cgaps@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 570 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 407.
FEATURES
Location/Qualifiers
1..437
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2365214"
/lab_hosts="DH10B"
/clone_lib="Soares NSF F8 9W OT PA P S1"
/notes="Organ: pooled; Vector: p773D-Pac (Pharmacia) with
a modified polylinker; Site: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clones: Soares NBHP pool 1:
30384-310919, 323208-325895 Soares NBHP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares NBHP pool 1:
758280-760583, 772104-774407 Soares NBHP pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NBHP
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 96 a 106 c 76 g 159 t
ORIGIN
Query Match 53.4%; Score 357.4; DB 9; Length 437;
Best Local Similarity 93.5%; Pred. No. 5.2e-83;
Matches 403; Conservative 0; Mismatches 6; Indels 22; Gaps 2;
QY 224 ATGAGTTGGCAGAGATCAACATATAGGCTTAGGCGCAGAGAGAGTTTACAGCTCTCT 283
DB 437 ATGAGTTGGCAGAGATC-ACCTATAGGCTTAGGCGCAGAGAGTTTACAGCTCTCT 379
QY 284 GAGCTGATTGGGGCTATGCTTACTGCTCCCTTGTCCAGGAGAACCCACTGATGAAGAG 343
DB 378 GAGCTGATTGGGGCTATGCTT-----GAGCCCACTGATGAAGAG 340
QY 344 CCTAAGAGAGAGAACCAACCCACTAAAGTCGGAATCTACCTGATCAGAGAGAGAA 403
DB 339 CCTAAGAA 280
QY 404 GATGATCAGGCTGAGCTGAGATTCAAGTGGCTCAGCTGAGGAGCGATCTCCAGGAGCTA 463
DB 279 GATGATCAGGCTGAGCTGAGATTCAAGTGGCTCAGCTGAGGAGCGATCTCCAGGAGCTA 220
QY 464 TGTGAGCAAAAGACTGGGGATGGATGTGAAGGTTGATGTGATGTCAAGGGGAGAGTTCTTA 523
DB 219 TGTGAGCAAAAGACTGGGGATGGATGTGAAGGTTGATGTGATGTCAAGGGGAGAGTTCTTA 160
QY 524 CCAAAGCAGAGAGCACTTTAAATGCGAAGCAGGTGAAGGGAATCACAGGTTTAAAGG 583
DB 159 CCAAAGCAGAGAGCACTTTAAATGCGAAGCAGGTGAAGGGAATCACAGGTTTAAAGG 100
QY 584 AAGATAAGCTGAACCAACACAACTGTTTTATATTAGATATTTTACTTTTAAAGAGCTTT 643
DB 99 AAGATAAGCTGAACCAACACAACTGTTTTATATTAGATATTTTACTTTTAAATATCTTT 40
QY 644 AATAAATTTT 654
|||||

Db 39 AATAAAGTTTT 29
RESULT 27
T87104
LOCUS
DEFINITION
IMAGE:115340 5', mRNA sequence.
T87104
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman
M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston
R., Williamson, A., Wohlmann, P. and Wilson, R.
The WashU-Merck EST Project
Unpublished
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 400
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: M13RP1
High quality sequence stop: 400.
FEATURES
Location/Qualifiers
1..555
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:470957"
/db_xref="taxon:9606"
/clone="IMAGE:115340"
/sex="male"
/dev_stages="20 week-post conception fetus"
/lab_hosts="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen LNPLS"
/note="Organ: liver and spleen; Vector: p773D (Pharmacia)
with a modified polylinker; Site: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGGAAGATTAAATTAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 147 a 126 c 152 g 127 t
ORIGIN
Query Match 52.6%; Score 351.8; DB 14; Length 555;
Best Local Similarity 87.6%; Pred. No. 1.6e-81;
Matches 496; Conservative 0; Mismatches 35; Indels 35; Gaps 9;
QY 14 CCGACTTCAGTGTGATGTTCTTGGACACCTGCTCAGTGTGATGTTCTCTGGGATC 73
DB 5 CAGCTTCAGTGTGATGTTCTTGGACACCTGCTCAGTGTGATGTTCTCTGGGATC 64
QY 74 TTCCCTTCGACCCCTTTCGCCACGTTGGAGCGCTGGGGAGCTGTGAGGTGTGAGGGGC 133
DB 65 TTCCCTTCGACCCCTTTCGCCACGTTGGAGCGCTGGGGAGCTGTGAGGTGTGAGGGGC 124
QY 134 AGTTTCCAGCGCTGTGAGATCTTTCTCTACTGAGACGAGCGCTATAGTCCGAGGC 193
DB 125 ACGTTCCAGCGCTGTGAGATCTTTCTCTACTGAGACGAGCGCTATAGTCCGAGGC 183

REFERENCE
AUTHORS

Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project

TITLE
Unpublished

JOURNAL
COMMENT

Contact: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Source: IMAGE Consortium, LNLN
High quality sequence stops: 337
This clone is available royalty-free through LNLN ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 663 Std Error: 0.00
Seq primer: Promega -21m13
High quality sequence stop: 337.

FEATURES
source

1..401
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="GBB:3771597"
/db_xref="taxon:9606"
/clone="IMAGE:202466"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/notes="Organ: Liver and Spleen; Vector: p77T3D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I oligo(dT) primer
[5' AACTGGAAGAATAAATTAAGACCTTTTTTTTTTTTTTT 3'] ,
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p77T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Patima Bonaldo."

BASE COUNT 85 a 98 c 72 g 141 t 5 others

Query Match 47.4%; Score 316.8; DB 14; Length 401;
Best Local Similarity 97.6%; Pred. No. 2.4e-72;
Matches 321; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 326 GAACCCCTCATGAGAGCCTTAAAGAGAGAACACCACCACTAAAAGTCGGATCCTACA 385
DB 336 GAACCCCTCATGAGAGCCTTAAAGAGAGAACACCACCACTAAAAGTCGGATCCTACA 277

QY 386 CCTGATCAGAGAGAGAAGATGATCATCGGTGCAGCTGAGATTCAAGTGCTGCACCTGGAA 445
DB 276 CCGTATCAGAGAGAGAAGATGATCATCGGTGCAGCTGAGATTCAAGTGCTGCACCTGGAA 217

QY 446 GCCGATCTCCAGGAGCTATGTCACAGAAAGACCTGGGATGATGATGAAGTGTGATCTGAT 505
DB 216 GCCGATCTCCAGGAGCTATGTCACAGAAAGACCTGGGATGATGATGAAGTGTGATCTGAT 157

QY 506 GTCAAGGGGAAGATTCTTACCAAAAAAGCAGACACTTTTAAATGCCAGAGCAGGTGAAGGG 565
DB 156 GTCAAGGGGAAGATTCTTACCAAAAAAGCAGACACTTTTAAATGCCAGAGCAGGTGAAGGG 97

QY 566 AAATCACAGGTTTAAAGGAAGATGATGTCACCAACCAAACTGTTTATATAGTATAT 625
DB 96 AAATCACAGGTTTAAAGGAAGATGATGTCACCAACCAAACTGTTTATATAGTATAT 37

QY 626 TTACTTTAAAGAGCTTTAAATAATTTTTT 654
DB 36 TTTAAAGTAAATAATCTTAAATAAGTTTTT 8

RESULT 32

W93171 389 bp mRNA linear EST 25-NOV-1999
zdr9f02.r1 Soares fetal heart_NbHH19W Homo sapiens cDNA clone
IMAGE:357051 5', mRNA sequence.

W93171 GI:1422324
EST.

ORGANISM
Homo sapiens (human)

REFERENCE
AUTHORS

Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project

TITLE
Unpublished

JOURNAL
COMMENT

Contact: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Source: IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 608 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 336.

FEATURES
source

1..389
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GBB:1273595"
/db_xref="taxon:9606"
/clone="IMAGE:357051"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal heart NBHH19W"
/notes="Organ: heart; Vector: p77T3D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAGTGGAGCGCCGACCTTTTTTTTTTTTTTT 3'] ,
double-stranded cDNA was size selected, ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p77T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Patima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBHL19K."

BASE COUNT 103 a 96 c 107 g 81 t 2 others

ORIGIN

Query Match 47.3%; Score 316.2; DB 14; Length 389;
Best Local Similarity 93.2%; Pred. No. 3.4e-72;
Matches 382; Conservative 4; Mismatches 24; Indels 4; Gaps 4;

QY 52 GTGTGCATGTTCTACTGGGCATCTTCCCTTCGACCCCTTTGCCACCTGGTGACCGCTGGG 111
DB 2 GTGTGCATGTTCTACTGGGCATCTTCCCTTCGACCCCTTTGCCACCTGGTGACCGCTGGG 61

QY 112 GAGCTGTGAGAGTGTGAGGGGCGACGTTTCCAGCGCTTGGACTCTTTCTCTCTCTACTGAGA 171
DB 62 GAGCTGTGAGAGTGTGAGGGGCGACGTTTCCAGCGCTTGGACTCTTTCTCTCTCTACTGAGA 121

QY 172 CGCAGCTTATAGTTCGCCAGSCAGTCTCTCCAGGAACTGAAATAGTGAATATGAGTTG 231
DB 122 CGCAGCTTATAGTTCGCCAGSCAGTCTCTCCAGGAACTGAAATAGTGAATATGAGTTG 181

QY 232 CGCAGGAAAGATCAACATATATAGGCTTAGGCCAAGAAGTTTTCACGCTCTCTGAGCTGAT 291

```

182 GCGAGGAGATCAATATAGGCTAGG-CAAGAGAGATTACAGCCTCTCTGAGTGAT 240
183 |||||
184 |||||
185 |||||
186 |||||
187 |||||
188 |||||
189 |||||
190 |||||
191 |||||
192 TGGGGCTATGCTTACTGGCTCCCTTTGTCGCCAGAACCCCTGATGAAGAGCCCTAAGA 351
193 |||||
194 |||||
195 |||||
196 |||||
197 |||||
198 |||||
199 |||||
200 |||||
201 |||||
202 |||||
203 |||||
204 |||||
205 |||||
206 |||||
207 |||||
208 |||||
209 |||||
210 |||||
211 |||||
212 |||||
213 |||||
214 |||||
215 |||||
216 |||||
217 |||||
218 |||||
219 TGGGGCTATGCTT-----GAACCCACTGATGAAGAGCCCTAAGA 279
220 |||||
221 |||||
222 |||||
223 |||||
224 |||||
225 |||||
226 |||||
227 |||||
228 |||||
229 |||||
230 |||||
231 |||||
232 |||||
233 |||||
234 |||||
235 |||||
236 |||||
237 |||||
238 |||||
239 |||||
240 |||||
241 |||||
242 |||||
243 |||||
244 |||||
245 |||||
246 |||||
247 |||||
248 |||||
249 |||||
250 |||||
251 |||||
252 |||||
253 |||||
254 |||||
255 |||||
256 |||||
257 |||||
258 |||||
259 |||||
260 |||||
261 |||||
262 |||||
263 |||||
264 |||||
265 |||||
266 |||||
267 |||||
268 |||||
269 |||||
270 |||||
271 |||||
272 |||||
273 |||||
274 |||||
275 |||||
276 |||||
277 |||||
278 |||||
279 |||||
280 |||||
281 |||||
282 |||||
283 |||||
284 |||||
285 |||||
286 |||||
287 |||||
288 |||||
289 |||||
290 |||||
291 |||||
292 |||||
293 |||||
294 |||||
295 |||||
296 |||||
297 |||||
298 |||||
299 |||||
300 |||||
301 |||||
302 |||||
303 |||||
304 |||||
305 |||||
306 |||||
307 |||||
308 |||||
309 |||||
310 |||||
311 |||||
312 |||||
313 |||||
314 |||||
315 |||||
316 |||||
317 |||||
318 |||||
319 |||||
320 |||||
321 |||||
322 |||||
323 |||||
324 |||||
325 |||||
326 |||||
327 |||||
328 |||||
329 |||||
330 |||||
331 |||||
332 |||||
333 |||||
334 |||||
335 |||||
336 |||||
337 |||||
338 |||||
339 |||||
340 GGGTGCAGTGAAGATTCAAGTGCCTGACCTGGGAAGCCGATCTCCAGGA 389
341 |||||
342 |||||
343 |||||
344 |||||
345 |||||
346 |||||
347 |||||
348 |||||
349 |||||
350 |||||
351 |||||
352 |||||
353 |||||
354 |||||
355 |||||
356 |||||
357 |||||
358 |||||
359 |||||
360 |||||
361 |||||
362 |||||
363 |||||
364 |||||
365 |||||
366 |||||
367 |||||
368 |||||
369 |||||
370 |||||
371 |||||
372 |||||
373 |||||
374 |||||
375 |||||
376 |||||
377 |||||
378 |||||
379 |||||
380 |||||
381 |||||
382 |||||
383 |||||
384 |||||
385 |||||
386 |||||
387 |||||
388 |||||
389 |||||
390 |||||
391 |||||
392 |||||
393 |||||
394 |||||
395 |||||
396 |||||
397 |||||
398 |||||
399 |||||
400 |||||
401 |||||
402 |||||
403 |||||
404 |||||
405 |||||
406 |||||
407 |||||
408 |||||
409 |||||
410 |||||
411 |||||
412 GGGTGCAGTG-AGATTCAAGTGCCTGACCT-GGAAGCCGATCTCCAGGA 459
413 |||||
414 |||||
415 |||||
416 |||||
417 |||||
418 |||||
419 |||||
420 |||||
421 |||||
422 |||||
423 |||||
424 |||||
425 |||||
426 |||||
427 |||||
428 |||||
429 |||||
430 |||||
431 |||||
432 |||||
433 |||||
434 |||||
435 |||||
436 |||||
437 |||||
438 |||||
439 |||||
440 |||||
441 |||||
442 |||||
443 |||||
444 |||||
445 |||||
446 |||||
447 |||||
448 |||||
449 |||||
450 |||||
451 |||||
452 |||||
453 |||||
454 |||||
455 |||||
456 |||||
457 |||||
458 |||||
459 |||||
460 |||||
461 |||||
462 |||||
463 |||||
464 |||||
465 |||||
466 |||||
467 |||||
468 |||||
469 |||||
470 |||||
471 |||||
472 |||||
473 |||||
474 |||||
475 |||||
476 |||||
477 |||||
478 |||||
479 |||||
480 |||||
481 |||||
482 |||||
483 |||||
484 |||||
485 |||||
486 |||||
487 |||||
488 |||||
489 |||||
490 |||||
491 |||||
492 |||||
493 |||||
494 |||||
495 |||||
496 |||||
497 |||||
498 |||||
499 |||||
500 |||||
501 |||||
502 |||||
503 |||||
504 |||||
505 |||||
506 |||||
507 |||||
508 |||||
509 |||||
510 |||||
511 |||||
512 |||||
513 |||||
514 |||||
515 |||||
516 |||||
517 |||||
518 |||||
519 |||||
520 |||||
521 |||||
522 |||||
523 |||||
524 |||||
525 |||||
526 |||||
527 |||||
528 |||||
529 |||||
530 |||||
531 |||||
532 |||||
533 |||||
534 |||||
535 |||||
536 |||||
537 |||||
538 |||||
539 |||||
540 |||||
541 |||||
542 |||||
543 |||||
544 |||||
545 |||||
546 |||||
547 |||||
548 |||||
549 |||||
550 |||||
551 |||||
552 |||||
553 |||||
554 |||||
555 |||||
556 |||||
557 |||||
558 |||||
559 |||||
560 |||||
561 |||||
562 |||||
563 |||||
564 |||||
565 |||||
566 |||||
567 |||||
568 |||||
569 |||||
570 |||||
571 |||||
572 |||||
573 |||||
574 |||||
575 |||||
576 |||||
577 |||||
578 |||||
579 |||||
580 |||||
581 |||||
582 |||||
583 |||||
584 |||||
585 |||||
586 |||||
587 |||||
588 |||||
589 |||||
590 |||||
591 |||||
592 |||||
593 |||||
594 |||||
595 |||||
596 |||||
597 |||||
598 |||||
599 |||||
600 |||||
601 |||||
602 |||||
603 |||||
604 |||||
605 |||||
606 |||||
607 |||||
608 |||||
609 |||||
610 |||||
611 |||||
612 |||||
613 |||||
614 |||||
615 |||||
616 |||||
617 |||||
618 |||||
619 |||||
620 |||||
621 |||||
622 |||||
623 |||||
624 |||||
625 |||||
626 |||||
627 |||||
628 |||||
629 |||||
630 |||||
631 |||||
632 |||||
633 |||||
634 |||||
635 |||||
636 |||||
637 |||||
638 |||||
639 |||||
640 |||||
641 |||||
642 |||||
643 |||||
644 |||||
645 |||||
646 |||||
647 |||||
648 |||||
649 |||||
650 |||||
651 |||||
652 |||||
653 |||||
654 |||||
655 |||||
656 |||||
657 |||||
658 |||||
659 |||||
660 |||||
661 |||||
662 |||||
663 |||||
664 |||||
665 |||||
666 |||||
667 |||||
668 |||||
669 |||||
670 |||||
671 |||||
672 |||||
673 |||||
674 |||||
675 |||||
676 |||||
677 |||||
678 |||||
679 |||||
680 |||||
681 |||||
682 |||||
683 |||||
684 |||||
685 |||||
686 |||||
687 |||||
688 |||||
689 |||||
690 |||||
691 |||||
692 |||||
693 |||||
694 |||||
695 |||||
696 |||||
697 |||||
698 |||||
699 |||||
700 |||||
701 |||||
702 |||||
703 |||||
704 |||||
705 |||||
706 |||||
707 |||||
708 |||||
709 |||||
710 |||||
711 |||||
712 |||||
713 |||||
714 |||||
715 |||||
716 |||||
717 |||||
718 |||||
719 |||||
720 |||||
721 |||||
722 |||||
723 |||||
724 |||||
725 |||||
726 |||||
727 |||||
728 |||||
729 |||||
730 |||||
731 |||||
732 |||||
733 |||||
734 |||||
735 |||||
736 |||||
737 |||||
738 |||||
739 |||||
740 |||||
741 |||||
742 |||||
743 |||||
744 |||||
745 |||||
746 |||||
747 |||||
748 |||||
749 |||||
750 |||||
751 |||||
752 |||||
753 |||||
754 |||||
755 |||||
756 |||||
757 |||||
758 |||||
759 |||||
760 |||||
761 |||||
762 |||||
763 |||||
764 |||||
765 |||||
766 |||||
767 |||||
768 |||||
769 |||||
770 |||||
771 |||||
772 |||||
773 |||||
774 |||||
775 |||||
776 |||||
777 |||||
778 |||||
779 |||||
780 |||||
781 |||||
782 |||||
783 |||||
784 |||||
785 |||||
786 |||||
787 |||||
788 |||||
789 |||||
790 |||||
791 |||||
792 |||||
793 |||||
794 |||||
795 |||||
796 |||||
797 |||||
798 |||||
799 |||||
800 |||||
801 |||||
802 |||||
803 |||||
804 |||||
805 |||||
806 |||||
807 |||||
808 |||||
809 |||||
810 |||||
811 |||||
812 |||||
813 |||||
814 |||||
815 |||||
816 |||||
817 |||||
818 |||||
819 |||||
820 |||||
821 |||||
822 |||||
823 |||||
824 |||||
825 |||||
826 |||||
827 |||||
828 |||||
829 |||||
830 |||||
831 |||||
832 |||||
833 |||||
834 |||||
835 |||||
836 |||||
837 |||||
838 |||||
839 |||||
840 |||||
841 |||||
842 |||||
843 |||||
844 |||||
845 |||||
846 |||||
847 |||||
848 |||||
849 |||||
850 |||||
851 |||||
852 |||||
853 |||||
854 |||||
855 |||||
856 |||||
857 |||||
858 |||||
859 |||||
860 |||||
861 |||||
862 |||||
863 |||||
864 |||||
865 |||||
866 |||||
867 |||||
868 |||||
869 |||||
870 |||||
871 |||||
872 |||||
873 |||||
874 |||||
875 |||||
876 |||||
877 |||||
878 |||||
879 |||||
880 |||||
881 |||||
882 |||||
883 |||||
884 |||||
885 |||||
886 |||||
887 |||||
888 |||||
889 |||||
890 |||||
891 |||||
892 |||||
893 |||||
894 |||||
895 |||||
896 |||||
897 |||||
898 |||||
899 |||||
900 |||||
901 |||||
902 |||||
903 |||||
904 |||||
905 |||||
906 |||||
907 |||||
908 |||||
909 |||||
910 |||||
911 |||||
912 |||||
913 |||||
914 |||||
915 |||||
916 |||||
917 |||||
918 |||||
919 |||||
920 |||||
921 |||||
922 |||||
923 |||||
924 |||||
925 |||||
926 |||||
927 |||||
928 |||||
929 |||||
930 |||||
931 |||||
932 |||||
933 |||||
934 |||||
935 |||||
936 |||||
937 |||||
938 |||||
939 |||||
940 |||||
941 |||||
942 |||||
943 |||||
944 |||||
945 |||||
946 |||||
947 |||||
948 |||||
949 |||||
950 |||||
951 |||||
952 |||||
953 |||||
954 |||||
955 |||||
956 |||||
957 |||||
958 |||||
959 |||||
960 |||||
961 |||||
962 |||||
963 |||||
964 |||||
965 |||||
966 |||||
967 |||||
968 |||||
969 |||||
970 |||||
971 |||||
972 |||||
973 |||||
974 |||||
975 |||||
976 |||||
977 |||||
978 |||||
979 |||||
980 |||||
981 |||||
982 |||||
983 |||||
984 |||||
985 |||||
986 |||||
987 |||||
988 |||||
989 |||||
990 |||||
991 |||||
992 |||||
993 |||||
994 |||||
995 |||||
996 |||||
997 |||||
998 |||||
999 |||||
1000 |||||

```

RESULT 33
N42010
LOCUS
DEFINITION
 Yw69f02.r1 Soares placenta 8to9weeks 2NBHP8to9W Homo sapiens CDNA
 clone IMAGE:257499 5', mRNA sequence.
ACCESSION
 N42010
VERSION
 N42010.1 GI:1166041
KEYWORDS
 EST.
SOURCE
 Homo sapiens (human)
ORGANISM
 Homo sapiens
REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
 Hillier, L., Clark, N., Dubucque, T., Elliston, K., Hawkins, M., Holman,
 M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
 Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston,
 R., Williamson, A., Wohlmann, P. and Wilson, R.
 The WashU-Merck EST Project
 Unpublished
TITLE
 Contact: Wilson RK
JOURNAL
 Washington University School of Medicine
COMMENT
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 High quality sequence stops: 321
 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Putative full length read
 Seq primer: T7
 High quality sequence stop: 321.
 Location/Qualifiers
 1..420
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:3887109"
 /db_xref="taxon:9606"
 /clone="IMAGE:257499"
 /dev_stage="two placentae: one from 8 weeks and another
 from 9 weeks post conception"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares placenta 8to9weeks 2NBHP8to9W"
 /notes="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dAT) primer [5'
 TGTTACCAATCTGAAGTGGAGCGCGCGGATTTTTTTTTTTTTTTT 3']
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT7T3 vector
 (Pharmacia). Library constructed by Bento Soares and
 M.Fatima Bonaldo."
BASE COUNT
 121 a 91 c 123 g 81 t 4 others
ORIGIN
 Query Match 46.8%; Score 313.4; DB 14; Length 420;
 Best Local Similarity 87.3%; Pred. No. 1.9e-71;

Matches 385; **Conservative** 0; **Mismatches** 30; **Indels** 26; **Gaps** 3;
QY 101 TCACCGCTGGGAGCTGTGAGAGTGTGAGGGGACAGTTCCAGCCGCTCTGGACTCTTTCTC 160
Db 1 TCACCGCTGGGAGCTGTGAGAGTGTGAGGGGACAGTTCCAGCCGCTCTGGACTCTTTCTC 60
QY 161 TCTTACTGAGACGCGACCTTATAGTTCGCGAGCCAGTCTCTCCAGGAACTGAAATAGTGA 220
Db 61 TCTTACTGAGACGCGACCTTATAGTTCGCGAGCCAGTCTCTCCAGGAACTGAAATAGTGA 120
QY 221 AATATGAGTTGGCGAGGAGATCAACATATAGGCTTAGGCCAAGAGAGTTTACAGCCT 280
Db 121 AATATGAGTTGGCGAGGAGATCAACATATAGGCTTAGGCCAAGAGAGTTTACAGCCT 180
QY 281 CTTGAGCTGATTGGGGCTATGCTTACTTGGCTCCCTTTGTTCCAGGAACTGATGAA 340
Db 181 CTTGAGCTGATTGGGGCTATGCTT-----GAACCCACTGATGAA 219
QY 341 GAGCCTAAAGAGAGAAACCCACCTAAAGTCCGAACTCTACACCTGATCAGAAGAGA 400
Db 220 GAGCCTAAAGAGAGAAACCCACCTAAAGTCCGAACTCTACACCTGATCAGAAGAGA 279
QY 401 GAAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCCAGGAG 460
Db 280 GAAGATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCCAGGAG 339
QY 461 CTATG-----TCAGACAAAGACTGGGATGGATGTGAAGT-GGTACTGATGCAAGGGGA 515
Db 340 CTATGTCAGACCAAGACTGGGGGATGGATGTGAAGTGGGTACTGATGTCAGGGGG 399
QY 516 AGATTCTACCAAAAGCAGAGC 536
Db 400 AAGNATTCCTANCAAAAAGC 420
RESULT 34
N40147
LOCUS
DEFINITION
 YW73e12.r1 Soares placenta 8to9weeks 2NBHP8to9W Homo sapiens CDNA
 clone IMAGE:257902 5', mRNA sequence.
ACCESSION
 N40147
VERSION
 N40147.1 GI:1163692
KEYWORDS
 EST.
SOURCE
 Homo sapiens (human)
ORGANISM
 Homo sapiens
REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
 Hillier, L., Clark, N., Dubucque, T., Elliston, K., Hawkins, M., Holman,
 M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
 Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston,
 R., Williamson, A., Wohlmann, P. and Wilson, R.
 The WashU-Merck EST Project
 Unpublished
TITLE
 Contact: Wilson RK
JOURNAL
 Washington University School of Medicine
COMMENT
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 High quality sequence stops: 450
 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: T7
 High quality sequence stop: 450.
 Location/Qualifiers
 1..547
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:3887512"
 /db_xref="taxon:9606"
 /clone="IMAGE:257902"
FEATURES
 source

/dev_stage="two placentae: one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares.placenta.8toWeeks.2NDHP8to9W"
/note="Organ: Placenta; Vector: pT73D (Pharmacia) with a
modified polylinker; site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
TGTTACCACTGATGAGGAGCGCGCGATTTTTTTTTTTT 3',
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fátima Bonaldo."
111 C 147 g 123 t 7 others
159 a

Query Match	45.6%	Score	304.8	DB	14	Length	547
Best Local Similarity	75.7%	Pred. No.	3.8e-69				
Matches	439	Conservative	0	Mismatches	58	Gaps	3
Qy	77	CTCTCGACCCCTTTCC	CCACCTGTGTGACCG	CTGGGAGCTGTGAGAGTGTGAGGGCACG	136		
Db	1	CCCTNGNCNATNTT	TCCACGTGTGAA	TCCCTCGAGTTGTGAGGGTGTGAGGGTCCG	60		
Qy	137	TTCCAGCCGTC	TGGACTCTTTCT	CTCTCTACTTGAGACGACGCTATAGGTC	CCGAGCCAG	196	
Db	61	TTCTCTGCTGTG	GGACTTTTTCT	GTCCACATGAGACGACGCTGT		104	
Qy	197	TCTCTCCAGGAAC	TGAAATAGTGA	AAATATAGTTGGCGAGGAAGATCA	ACATATAGGCT	256	
Db	105	-----GTG	AAATATGATTTT	GGCGAGGAAGATCA	CAATATAGGCT	144	
Qy	257	AGCCCAAGAAG	TTTACAGCTCT	CTGAGCTGATTTGGGCTATGCTT	ACTGCTGCCCT	316	
Db	145	AGCGGAAGAG	AGTGTACCA	CTCTGAGCTGATTTGGGCTATGCT		191	
Qy	317	TTGTCCAGGA	ACCCACTGTATG	AGAGGCTTAAAGAGAGAA	CCACCCACTAAAGTCGG	376	
Db	192	-----GGAG	CCCGGTGATGAG	GAGGCTCAGCAAGAGGA	CCACCACTGAAAGTCGG	243	
Qy	377	AATCCTACAC	CTTGATCAGA	AAAGAGAAAGATGATCAGG	GTGCAGCTGAGATTCAAGTCCT	436	
Db	244	GATCCTGCAC	CTGGTCAGG	AGAGAGAGAAAGATCAGG	GTGCAGCTCAAGTCCT	303	
Qy	437	GACTGGAAG	CCGATCTCCAG	AGCTATGTTCAGACAA	GACTGGGATGATCTGAAGCT	496	
Db	304	GACTGGAAG	CTGATCTCAG	AGAGCTGTCTCAGTCAA	AGACTGTGGGTGTAATGTGAAT	363	
Qy	497	GGTACTGAT	CTCAAGGGG	GAAGATTCTACCA	AAAGCAGACGACTTTTAAATGCCAGAAGCA	556	
Db	364	GGTCTGTAT	GCAC	CGGGGAAGATTCTGCC	AAAAATCAGAA	CAATTTTAAATATGCCAGAAGCA	423
Qy	557	GGTCAAGG	GAATTCAG	CTTTTAAAGG	AGATATAGCTGAAACACAC	CAACAACTGTTTTT	615
Db	424	GGTGAC	CAGGCA	CCACAGGTTTAAAT	TAAGACAGCTGTAAC	CAACACAACTGTTTTTA	483
Qy	616	TATPAGATAT	TTTACTTTAA	GAGTCCTTA	TAAATTTTTTG	655	
Db	484	TCTAAGATAT	TTTCACTTNA	AAATATATCG	ANATTAACCTTTTG	523	

RESULT 35
R99925/c
LOCUS
DEFINITION
R99925 410 bp mRNA linear EST 14-SEP-1995
Yq76f02.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:201723 3', mRNA sequence.
ACCESSION R99925
VERSION R99925
KEYWORDS R99925.1 GI:986526
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE	AUTHORS	TITLE	JOURNAL
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

BASE COUNT	91 a	98 c	75 g	140 t	6 others
ORIGIN					
Query Match	45.3%	Score 303.2;	DB 14;	Length 410;	
Best Local Similarity	89.7%;	Pred. No. 9,1e-69;			
Matches 376;	Conservative	0;	Mismatches 18;	Indels 25;	Gaps 47;
Qy	240	GATCAACATATAGGCTTAGGCCAAGAAGAGTTTACAGCCTCTCTGA--GCTGATTGGGGC	297		
Db	410	GATCAACATATAGGCCCTAGGCCAAGAANAAGTTTACAGCCCTCCTGAGNCTGATTGGGGC	351		
Qy	298	TATGCTTATCGGCTCCCTTTGTGCCAGAACCCACTGATGAAGG--CCTTAACAGAGAGA	356		
Db	350	TATGCTT-----GAAACCCTGNTGAAGAGCCCTTAAGNAGAGA	312		
Qy	357	AACCAACCCTAAAAGTCGGAATCTTACACCTGATCAGAAAGAGAGAAGATGATCAGGGTG	416		
Db	311	AACCAACCCTAAAAGTCGGAATCTTACACCTGATCAGAAAGAGAGAAGATGATCAGGGTG	252		
Qy	417	CAGCTGAGATTCAAGTGCCTGA--CCTGGAGCCGATCTCCAGGAGCTATGTGCACACAAAG	475		
Db	251	CAGCTGAGATTCAAGTGCCTGAGNCTTGAAGCCGATCTCCAGGAGCTATGTGCACACAAAG	192		
Qy	476	ACTGGGGATGGATGTGAAGTGGTACTGTGTCAAGGGGAAGATTCTACCAAAAAGCGAG	535		
Db	191	ACTGGGGATGGATGTGAAGTGGTACTGTGTCAAGGGGAAGATTCTACCAAAAAGCGAG	132		
Qy	536	CATTTTAAATGCCAGAGCAGGTGAAGGAAATACACAGGTTTAAAGCAAGATAAGCTGA	595		
Db	131	CNCTTTAAATGCCAGAGCAGGTGAAGGAAATACACAGGTTTAAAGCAAGATAAGCTGA	72		
Qy	596	AACAACACAACCTGTTTTTATATPAGATATTTTACTTTTAAAGCTCTTATATAATTTTT	654		

```

71 AACACACAACTGTTTATATTAGATATTTTAAAGTAAAAATCTTATTAAGTTTT 13
|||||
Db
RESULT 36
CB992682
LOCUS
DEFINITION
AGENCOURT 13633348 NIH_MGC_148 Homo sapiens cDNA clone
IMAGE:30338928 5', mRNA sequence.
ACCESSION
CB992682
VERSION
CB992682.1 GI:30287202
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 570)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM366 row: g column: 01
High quality sequence stop: 558.
Location/Qualifiers
1. 570
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30338928"
/tissue_type="pre-eclamptic placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_148"
/note="Organ: Placenta; Vector: pBluescriptR; Site:1:
all-XhoI; Site:2: BamH; Library is oligo-dT primed and
directionally cloned using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert
size 2.3 kb and normalized to ROT 5. This is a primary
library enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation).
Library constructed by M. Brownstein (NIH/NHGRI).
National Institutes of Health). Note: this is a NIH_MGC
Library."
BASE COUNT 174 a 116 c 157 g 123 t
ORIGIN
Query Match 44.3%; Score 296.4; DB 14; Length 570;
Best Local Similarity 76.2%; Pred. No. 6.2e-67;
Matches 430; Conservative 0; Mismatches 76; Indels 58; Gaps 3;
QY 93 CCACGTGTGACCGCTGGGAGCTGTGAGAGTGTGAGGGGACGTTCCAGCCGCTGGAC 152
Db 24 CCAGGCTGTGAATGCCCTGGAGTGTGAGGGTGTGAGGGTCCGCTTCCTGCTGTGGAC 83
QY 153 TCTTTCTCTCTACTGAGAGCGAGCTATAGTCCGCGAGCCAGTCTCCAGGACTGA 212
Db 84 TTTTCTGTCCCACTGAGACGAGCTGT----- 111
QY 213 AATAGTGAATATGAGTTGGCGAGGAAGATCAACATATAGGCTAGGCCAAGAAGATT 272
Db 112 ----GTGAATATGATTTGGCGAGGAAGATCAACATATAGGCTAGGCCGAGGAGAGTG 167
QY 273 TACAGCTCTGAGCTGATGGGCTATGCTTACTGGCTCCCTTTGTCACGAGNACCA 332
Db 168 TACCACCTCCTGAGCTGATGGGCTATGCT-----GGAGCCCG 206

```

```

333 CTGATGAAGAGCCCTAAAGAGAGAAACCCACCTAAAGTCGGAATCCTACACCTGATC 392
|||||
207 GTGATGAGGAGCCCTCAGCAAGAGGAAACCCACCACTGAAAGTCGGGATCTCTCACCTGTC 266
|||||
393 AGAAGAGAGAGATGATCAGCGGTGACCTGAGATTCAGTGCCTGACCTGGAAGCCGATC 452
|||||
267 AGGAGAGAGAGAGATCAGCGGTGACCTGAGATTCAGTGCCTGACCTGGAAGCTGATC 326
|||||
453 TCCAGGAGCTATGTGAGCAAAAGACTGGGGATGGATGTGAAAGGTGCTGATGTCAGG 512
|||||
327 TCCAGGAGCTGTCTCAGTCAAGACTGGGGTGAATGTGGAATGGTCTCTGATGACCAAG 386
|||||
513 GGAAGATTCACCAAGACGAGCATTAAATGCCAGAGCAGGTGAAGGAATCAC 572
|||||
387 GGAAGATTCGCCAAATCGAACAATTTAAATGCCAGAGGAGGTGACAGCAACAC 446
|||||
573 AGGTTAAAGAGAGATAAGCTGAAACACAC-AAAAGCTTTTATATTAGATATTTTACT 631
|||||
447 AGGTTAAATGAGACAGCTGAAACACACAACTGTTTTTATTAAGATATTGACT 506
|||||
632 TTAAGAGCTCTTAATAAATTTTG 655
|||||
507 TAAATATCAATAAATTTTG 530
|||||
CB994958 570 bp mRNA linear EST 01-MAY-2003
AGENCOURT 13643389 NIH_MGC_148 Homo sapiens cDNA clone
IMAGE:30335086 5', mRNA sequence.
CB994958
VERSION
CB994958.1 GI:30289478
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 570)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM356 row: f column: 23
High quality sequence stop: 570.
Location/Qualifiers
1. 570
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30335086"
/tissue_type="pre-eclamptic placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_148"
/note="Organ: Placenta; Vector: pBluescriptR; Site:1:
all-XhoI; Site:2: BamH; Library is oligo-dT primed and
directionally cloned using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert
size 2.3 kb and normalized to ROT 5. This is a primary
library enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation).
Library constructed by M. Brownstein (NIH/NHGRI).
National Institutes of Health). Note: this is a NIH_MGC
Library."
FEATURES
source

```



```
BASE COUNT      175 a      115 c      157 g      123 t
ORIGIN
Query Match      44.3%; Score 296.4; DB 14; Length 570;
Best Local Similarity 76.2%; Pred. No. 6.2e-67;
Matches 430; Conservative 0; Mismatches 76; Indels 58; Gaps 3;
QY 93 CCAGGTGTCACCGCTGGGAGCTGTGAGAGTGTGAGGGGCACTGTCACCGCTCTGGAC 152
Db 24 CCAGGTGTCACCGCTGGGAGCTGTGAGAGTGTGAGGGGCGGCTCTCTGGAC 83
QY 153 TCTTTCTCTCTACTAGAGCGCAGCTTATAGTCCGCGAGCCAGTCTCTCCAGGAAGTGA 212
Db 84 TTTTCTGTCCACTGAGAGCGAGCTGT----- 111
QY 213 AATAGTGAATATAGTGTGGGAGGAGATCAACATATAGGCTAGGCCAAGAGAGTT 272
Db 112 ----GTGAATATAGTGTGGGAGGAGATCAACATATAGGCTAGGCCAAGAGAGTG 167
QY 273 TACAGCTCTCTGAGCTGATTTGGGGCTATGCTTACTGCTCCCTTTGTCCAGGAACCCA 332
Db 168 TACACCTCTGAGCTGATTTGGGCTATGCT-----GGAGCCCG 206
QY 333 CTGATGAAGAGCTTAAAGAGAGAGAACACCGACTAAAGTCGAATCTTACACTGATC 392
Db 207 GTGATGAGGAGCTTCAGCAAGAGAGAACCACTGAAGTCGGGATCTCTGCACTGGTC 266
QY 393 AGAAGAGAGAGAGATGATCAGGCTGACGTGAGATTCAGTGCCTGAGCGCGATC 452
Db 267 AGGAGAGAGAGAGATCAGGCTGACGTGAGATTCAGTGCCTGAGCGCGATC 326
QY 453 TCCAGGAGCTATGTCAGCAAAAGACTGGGATGATGTAAGTGATGTAAGTGTAAGG 512
Db 327 TCCAGGAGCTGTCTCAGTCAAGAGACTGGGGTGTAATGTGGAATGTCTCTGATGAC 386
QY 513 GGAAGATCTTACCAAGAGAGAGACACTTAAATGCCAGAGAGAGGTAAGGGAATCAC 572
Db 387 GGAAGATCTTCCCAATTCAGAACATTTAAATGCCAGAGAGAGGTAAGGGAATCAC 446
QY 573 AGGTTTAAAGAGAGATAGCTGAACCAACAC-AAACTGTTTTTATATAGATATTTACT 631
Db 447 AGGTTTAAATGAGACAGCTGAACCAACCAACAACTGTTTTTATCTAAGATATTTGACT 506
QY 632 TTAAGAGCTCTTAAATATTTTG 655
Db 507 TAAATAATCAAAATAAACTTTTG 530

RESULT 38
BG354572      499 bp      mRNA      linear      EST 05-MAR-2001
LOCUS      BG354572
DEFINITION      PLAC6 Human placenta cDNA expression library Homo sapiens CDNA
clone 253708 similar to PLAC6 Placenta specific gene 6, mRNA
sequence.
ACCESSION      BG354572
VERSION      BG354572.1 GI:13198770
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      1 (bases 1 to 499)
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE      Placental genes identified by large-scale expression analysis
JOURNAL      Unpublished
COMMENT      Contact: Michael Walker
Incyte Genomics
3160 Porter Drive, Palo Alto, CA 94304, USA
Tel: 650 845-5771
Fax: 650 621-8514
Email: mwalker@incyte.com.
Location/Qualifiers
1..499

FEATURES
Source
```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="253708"
/clone_lib="Human placenta cDNA expression library"
/notes="Organ: placenta"
BASE COUNT      148 a      100 c      139 g      112 t
ORIGIN
Query Match      43.4%; Score 290.2; DB 10; Length 499;
Best Local Similarity 76.9%; Pred. No. 2.5e-65;
Matches 419; Conservative 0; Mismatches 68; Indels 58; Gaps 3;
QY 112 GAGCTGTGAGAGTGTGAGGGGCACTTCCAGCGCTGTGGACTCTTCTCTCTCTACTGAGA 171
Db 1 GAGTTGTGAGGGTGTGAGGGTGTGAGGGTGTGAGGGTGTGAGGGTGTGAGGGTGTGAG 60
QY 172 CGAGCTTATAGTGTGCGGAGCGGCACTCTCCAGGAAGTGAATAGTGAATATAGTTG 231
Db 61 CGAGCTGT-----GTGAATATAGTTG 84
QY 232 CGGAGAGAGATCAACATATAGGCTAGGCCAAGAGAGTTTACAGCTCTCTGAGCTGAT 291
Db 85 GCGAGAGAGATCAACATATAGGCTAGGCCAAGAGAGTGTACACCTCTCTGAGCTGAT 144
QY 292 TGGGCTATGCTTACTGCTCCCTTTGTCCAGGAACCCCACTGATGAAGAGCCCTAAAGA 351
Db 145 TGGGCTATGCT-----GGAGCCGCTGATGAGGAGCCCTCAGCA 183
QY 352 AGAGAAACACCCACTAAAGTCGGAATCCTACACCTGATCAGAGAGAGAGATGATCA 411
Db 184 AGAGAAACACCCACTAAAGTCGGAATCCTGACCTGTCAGGAGAGAGAGATGATCA 243
QY 412 GGGTCAGCTGAGATTCAGTGCCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 471
Db 244 GGGTCAGCTGAGACTCAAGTGCCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 303
QY 472 AAAGCTGGGATGATGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 531
Db 304 AAAGCTGGGATGATGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 363
QY 532 AGAGCACTTTAAATGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 591
Db 364 AGAACAATTTAAATGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
QY 592 CTGAACAACAC-AAACTGTTTTTATATAGATATTTTACTTTTAAAGAGTCTTAAAT 650
Db 424 CTGAACAACACAAAACTGTTTTTATCTAAGATATTTGACTTAAATAATCAAAATAAAC 483
QY 651 TTTTG 655
Db 484 TTTTG 488

RESULT 39
CB961487      520 bp      mRNA      linear      EST 29-APR-2003
LOCUS      CB961487
DEFINITION      AGENCOURT 13894159 NIH.MGC.148 Homo sapiens cDNA clone
IMAGE:30348215 5', mRNA sequence.
ACCESSION      CB961487
VERSION      CB961487.1 GI:30217604
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      1 (bases 1 to 520)
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE      NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
```



```

RESULT 41
AI742551/c
LOCUS
DEFINITION
  AI742551 503 bp mRNA linear EST 19-DEC-1999
  w55c11.x1 Soares NSF F8 9W OT PA P S1 Homo sapiens cDNA clone
  IMAGE:2369012 3' similar to SW:G022_HUMAN Q13066 GAGE-2 PROTEIN. [1
  1 //, mRNA sequence.
ACCESSION
AI742551 1 GI:5110839
VERSION
AI742551.1
KEYWORDS
Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 503)
REFERENCE
  NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-remail.nih.gov
  This clone is available royalty-free through LLNL; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Insert Length: 1035 Std Error: 0.00
  Seq primer: -40UP from Gibco
  High quality sequence stop: 454.
FEATURES
  source
    1..503
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:2369012"
    /lab_host="DH10B"
    /clone_lib="Soares NSF F8 9W OT PA P S1"
    /notes="Organ: pooled; Vector: pT730-Pac (Pharmacia) with
    a modified polylinker; Site 1: Not I; Site 2: Eco RI;
    Equal amounts of plasmid DNA from five normalized
    libraries were mixed, and ss circles were made in vitro.
    Following HAP purification, this DNA was used as tracer
    a subtractive hybridization reaction. The driver was
    PCR-amplified cDNAs from pools of 5,000 clones made from
    the same 5 libraries. The pools consisted of the following
    libraries and clones: Soares NbHSP pool 1:
    30384-310919, 323208-325895 Soares Nb2HP pool 1:
    145032-147335, 147720-148103, 148872-149255, 15002 -
    150407, 151176-152327 Soares Nb2HF8-9W pool 1:
    758280-760583, 772104-774407 Soares NbHFA pool 1:
    304776-306311, 320136-322823, 326280-326663 Soares NbHOT
    pool 1: 723720-726407, 739080-740999 Subtraction by Bento
    Soares and M. Fatima Ronaldo."
    105 a 121 c 98 g 179 t
BASE COUNT
  105 a 121 c 98 g 179 t
Query Match 42.0%; Score 280.8; DB 9; Length 503;
Best Local Similarity 81.1%; Pred. No. 7.6e-63;
Matches 360; Conservative 0; Mismatches 62; Indels 22; Gaps 2;

Qy 213 AATAGTCAATATACAGTGGGAGGAGATCAATATAGCTAGCCAGGAGAGT 272
Db 455 AATGTGAATATGATTGGCGAGAGATCAATATAGCTAGCCAGGAGAGT 396
Qy 273 TACAGCTCTGAGCTGATGGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAACCCA 332
Db 395 TACCACCTCTGAGCTGATGGGCTATGCTT-----GGAGCCG 357
Qy 333 CTGATGAAGACCTTAAGAGAGAGAACCCACCTAAAGTCGATCTTACCTGATC 392
Db 356 GTGATGAGGAGCCTCAGCAGAGAGAACCCACCACTGAAAGTCGGATCTCGACCTGGTC 297
Qy 393 AGAAGAGAGAGATGATCAGGCTGAGTCAAGTTCAGTGCCTGAGCTGGAAGCCGATC 452
Db 296 AGGAGAGAGAGAGATCAGGGTTCACTAGAGATCAAGTGCTGACCTGGAAGCTGATC 237

```

```

Qy 453 TCCAGGAGCTATGTCAGACAAAGACTGGGATCGATGTGAGCTGCTGACTGTCAAGG 512
Db 236 TCCAGGAGCTGTCTCAGTCAAGAGACTGGGGTGAAATGTTGGAATGCTCTGATGACGAG 177
Qy 513 GGAAGATTCTACCAAAAGACAGACACTTTTAAATGCCAGAGCAGGTGAAGGAAATCAC 572
Db 176 GGAAGATTCTCCCAAAATCAGAACAAATTTAAATGCCAGAGCAGGTGACAGCAACCAC 117
Qy 573 AGTTTAAAGAGAGATGAGCTGAACACAC-AAACTGTTTTTATTATAGATTTTACT 631
Db 116 AGTTTAAATGAAGACAAAGCTGAACAAACCCAAACTGTTTTTATTATAGATTTTACT 57
Qy 632 TTAAGAGTCTTAAATAAATTTTG 655
Db 56 TAAATAATCGAATAAATTTTG 33
RESULT 42
BG436305
LOCUS
DEFINITION
  BG436305 849 bp mRNA linear EST 14-MAR-2001
  602508742P1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4605658 5',
  mRNA sequence.
ACCESSION
BG436305
VERSION
BG436305.1 GI:13342811
KEYWORDS
EST.
SOURCE
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 849)
REFERENCE
  NIH-MGC http://mgs.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-remail.nih.gov
  Tissue Procurement: CLONTECH Laboratories, Inc.
  cDNA Library Preparation: CLONTECH Laboratories, Inc.
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA sequencing by: Incyte Genomics, Inc.
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  plate: LLCM1343 row: d column: 11
  High quality sequence stop: 514.
FEATURES
  source
    1..849
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:4605658"
    /lab_host="DH10B (T1 phage-resistant)"
    /clone_lib="NIH_MGC_79"
    /notes="Organ: placenta; Vector: pDNR-LIB (Clontech);
    Site 1: SfiI (ggcgctcgcc); Site 2: SfiI (ggcgctatggcc
    ); 5' and 3' adaptors were used in cloning as follows: 5'
    adaptor sequence: 5'-CAGGCCATTATGGC-3' and 3' adaptor
    sequence: 5'-ATTCTAGAGCGCGGCGGACATG-dt(30)BN-3'
    (where B = A, C, or G and N = A, C, G, or T). Average
    insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies
    contained inserts by PCR. This library was enriched for
    full-length clones and was constructed by Clontech
    Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
    Library."
    263 a 177 c 262 g 147 t
BASE COUNT
  263 a 177 c 262 g 147 t
ORIGIN
  41.6%; Score 278.6; DB 10; Length 849;
  Best Local Similarity 74.2%; Pred. No. 3.4e-62;
  Matches 406; Conservative 0; Mismatches 84; Indels 57; Gaps 2;

Qy 102 GACCCCTGGAGCTGTGAGAGTGTGAGGGGACCTTCCAGCCGCTCGACTCTTCTCT 161

```

Db 2 GAATGCCCTGGAGTTGTGAGGGTGTGAGGGTGGCGTTCTCTGCTCTCTGCACTTTTCTGT 61
 QY 162 CCTACTGAGAGCGCAGCCTATAGGTCGCGCAGGCCAGTCTCTCCAGGAAGTGAAGTGA 221
 Db 62 CCCACTGAGAGCGCAGCTGT-----GTGAA 85
 QY 222 ATATGAGTTGGGAGGAGATCAACATATAGGCTAGGCTAGGCCAAGAGAGTTTACAGCTC 281
 Db 96 ATATGATTGGGAGGAGATCAACATATAGGCTAGGCTAGGCCAAGAGAGTTTACAGCTC 145
 QY 282 CTGAGCTGATTGGGGCTATGCTTACTGCTCCCTTTGCTCCAGGAACCCCACTGATGAAG 341
 Db 146 CTGAGCTGATTGGGGCTATGCT-----GGAGCCCGGTGATGAGG 184
 QY 342 AGCCTAAGAGAGAGAAACACCCCTAAAGTGGGATCTTACCTGATCAGAGAGAG 401
 Db 185 AGCCTCAGCAGAGAGAGAACCAACTGAAAGTGGGATCTTCCGATCGGTCAGAGAGAG 244
 QY 402 AAGATGATCAGGTCGAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCCAGAGC 461
 Db 245 AAGAGATCAGGTCGAGCTGAGATTCAAGTGCCTGACCTGGAAGCTGATCTCCAGAGC 304
 QY 462 TATGTCAGCAAGAGCTGGGATGATGTGAAGTGGTACTGATGTCAAGGGGAGATTC 521
 Db 305 TGTCTCAGTCAAGAGCTGGGGTGAATGTGGAATGGTCTGATGACCCAGGGGAGATTC 364
 QY 522 TACCAAGAGCAGAGCAGCTTTAAATGCGCAGAGCAGGGTGAAGGAATCAAGGTTTAA 581
 Db 365 TGCCAAATCAGAACATTTAAATGCCAGAGAGGGTGAAGGAGTCAAGCAACCAAGTTTAA 424
 QY 582 GGAAGATAGCTGAAACCAACAACTGTTTTATATTAGATATTTTACTTTAAAGAGTC 641
 Db 425 TGAAGACAGCTGAAACCAACAACTGTTTTATCTAAGATGTTTGACTTAAAGAAAT 484
 QY 642 TTAATAA 648
 Db 485 ATCAAA 491

RESULT 43
 A1189067/c
 LOCUS
 DEFINITION
 qd18a09, xl Soares placenta, 8to9weeks, 2NBHP8to9W Homo sapiens cDNA
 clone IMAGE:1724056 3', similar to SW7GGE2_HUMAN Q13066 GAGE-2
 PROTEIN. [1] ; mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 429)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov

This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 602 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 363.
 Location/Qualifiers

FEATURES
 source
 1..429
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1724056"
 /dev_stage="two placenta: one from 8 weeks and another
 from 9 weeks post conception"
 /lab_host="DH10B (ampicillin resistant)"

BASE COUNT 87 a 115 c 85 g 142 t
 ORIGIN

Query Match 41.5%; Score 277.8; DB 9; Length 429;
 Best Local Similarity 81.8%; Pred. No. 4.4e-62;
 Matches 354; Conservative 0; Mismatches 57; Indels 22; Gaps 2;

QY 224 ATGAGTTGGCGAGGAGATCAACATATAGGCTAGGCCAAGAGAGTTTACAGCTCTCT 283
 Db 429 ATGATTGGCGAGGAGATCAACATATAGGCTAGGCCGAGGAGAGGGTACCACTCTCT 370
 QY 284 GAGCTGATTGGGGCTATGCTTACTGCTCCCTTTGTCCTCCAGGAACCCCACTGATGAAGAG 343
 Db 369 GAGCTGATTGGGGCTATGCT-----GGAGCCCGGTGATGAGGAG 331
 QY 344 CCTAAGAGAGAGAACCCCACTAAAGTGGGATCTTACACCTGATCAGAGAGAGAA 403
 Db 330 CCTCAGAGAGAGAACCCCACTAAAGTGGGATCTTACACCTGATCAGAGAGAGAA 271
 QY 404 GATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAGCCGATCTCCAGAGCTA 463
 Db 270 GAAGATCAGGGTGCAGCTGAGACTCAAGTGCCTGACCTGGAAGCTGATCTCCAGAGCTG 211
 QY 464 TGTCTCAGCAAGAGCTGGGGATGGATGCTGAGTGTGAGTGTCTCAAGGGGAGATTTCTA 523
 Db 210 TCTCAGTCAAGAGCTGGGGTGAATGTGGAATGGTCTGATGACAGGGGAGATTTCTG 151
 QY 524 CCAAGAGCAGAGCACTTTAAATGCCAGAGCAGGTGAAGGGAAATCAACAGTTTAAAGG 583
 Db 150 CCAAAATCAGAACAAATTTAAATGCCAGAGCAGGTGACAGCAACCAACAGTTTAAATG 91
 QY 584 AGATGAGCTGAAACAAACAC-BACTGTTTTTATTATAGATATTTTACTTTAAAGAGTCT 642
 Db 90 AAGCAAGCTGAAACAAACAAACAACTGTTTTTATCTAAGATATTTGACTTAAATAATCG 31
 QY 643 TAATAAATTTTGTG 655
 Db 30 AAATAAATTTTG 18

RESULT 44
 H87913
 LOCUS

DEFINITION
 yw1608.r1 Soares placenta, 8to9weeks, 2NBHP8to9W Homo sapiens cDNA
 clone IMAGE:252446 5', mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 544)
 Hillier, L., Clark, N., Dubuque, T., Eliston, K., Hawkins, M., Holman,
 M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
 Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston,
 R., Williamson, A., Wohldmann, P. and Wilson, R.

The WashU-Merck EST Project
 Unpublished
 Contact: Wilton RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800

/clone lib="Soares placenta 8to9weeks 2NBHP8to9W"
 /note="Organ: placenta; Vector: pT73D (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5',
 TGTTCACCAATCTGAAGTGGAGCGCGCGGATTTTTTTTTTTTTTTT 3',],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT73 vector
 (Pharmacia). Library constructed by Bento Soares and
 M. Fatima Bonalogo."

Fax: 314 286 1810
 Email: est@watson.wustl.edu
 High quality sequence stops: 366
 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 663 Std Error: 0.00
 Seq primer: M13Rp1
 High quality sequence stop: 366.

FEATURES

Location/Qualifiers
 1..544
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:3885512"
 /db_xref="taxon:9606"
 /clone="IMAGE:252446"
 /dev_stage="two placentae: one from 8 weeks and another
 from 9 weeks post conception"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares Placenta 8to9weeks 2NHP8to9W"
 /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5',
 TGTTACCAACTGAAGTGGGAGCGCGCGATTTTTTTTTTTT 3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT7T3 vector
 (Pharmacia). Library constructed by Bento Soares and
 M.Fatima Bonaldo."
 127 a 121 c 149 g 132 t 15 others

BASE COUNT

127 a 121 c 149 g 132 t 15 others

ORIGIN

Query Match 39.9%; Score 266.8; DB 14; Length 544;
 Best Local Similarity 81.3%; Pred. No. 3.7e-59;
 Matches 427; Conservative 0; Mismatches 59; Indels 39; Gaps 9;
 QY 31 GTTCCTTGACACCTGCTCAGTGTGATGTCATGGGATCTTCCTTCGACCCCTTT 90
 Db 43 GCTCCTTGACACCTGCTCAGTGTGATGTCATGGGATCTTCCTTCGACCCCTTT 102
 QY 91 GCCACGTGTGACCGCTGGGAGCTGTGAGAGTGTGAGGGGACAGTTCACGCGTCTGG 150
 Db 103 GCCACGTGTGACCGCTGGGAGCTGTGAGAGTGTGAGGGGACAGTTCACGCGTCTGG 162
 QY 151 ACTCTTCTCTCTACTGACGAGCAGCTATAGTTCGCGAGCGCAGTCTCCAGGAACT 210
 Db 163 ACTCTTCTCTCTACTGACGAGCAGCTATAGTTCGCGA-GCCAGTCTCTCCAGGAACT 221
 QY 211 GAAATAGTGAATATGAGTTGGCGAGGAAGATCAACATATAGGCTAGGCGGCAAGAG 270
 Db 222 GAAATAGTGAATATGAGTTGGCGAGGAAGATCAACATATAGGCTAGGCGGCAAGAG 280
 QY 271 TTTACAGCTCTGAGCTGATGGGGCTATGCTTACTGCTCCCTTTGTCCCGAGAAC 330
 Db 281 TTTACAGCTCTGAGCTGATGGGGCTATGCTT-----GAACT 319.
 QY 331 CACTGATGAAGAGCCTAAAGAGAAACACCCACCTAAAGTCCGAA--TCCTACACCT 388
 Db 320 CACTGATGAAGAGCCTAAAGAGAAACACCCACCTAAAGTCCGAA--TCCTACACCT 379
 QY 389 G-----ATCAGNAGAGAGATGATCAGGGTGCAGCTGAGATTCAA---GTGCTGACC 440
 Db 380 GAAATCAGNAGAGAGAGATGATCAGGGTGCAGCTGAGATTCAAAGTTCGCTGAACTN 439
 QY 441 TGGAGAGCGATCTCCAGGAGCTAT--GTTCAGACAAAGACTCGGGATGG---ATGTGAAGG 495
 Db 440 GGGAGAGCGTCTTCAGGAGCTTTTGTTCAGACAAAGATGGGNTTGGATGTTGNAAGG 499
 QY 496 TGGTACTGATG--TGAGGGGAGAGATTCTACCAAGAGCAGCACT 539
 Db 500 TTGTATTGATGTTCAAGGGGAGTTTTCCTCCAAAGAGGAGGCGT 544

RESULT 45

N27023/c
 LOCUS
 DEFINITION
 yw73el2.s1 Soares placenta_8to9weeks_2NHP8to9W Homo sapiens cDNA
 clone IMAGE:257902 3', mRNA sequence.
 N27023
 VERSION
 N27023.1 GI:1141371
 EST.
 KEYWORDS
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1. (bases 1 to 430)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman,
 M., Lultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
 Rifkin, L., Rohlfing, T., Soares, M., Tan, P., Trevasakis, E., Waterston,
 R., Williamson, A., Wohlmann, P. and Wilson, R.
 The WashU-Merck EST Project
 Unpublished
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 High quality sequence stops: 296
 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: m13 -40 forward
 High quality sequence stop: 296.
 Location/Qualifiers
 1..430
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:3887512"
 /db_xref="taxon:9606"
 /clone="IMAGE:257902"
 /dev_stage="two placentae: one from 8 weeks and another
 from 9 weeks post conception"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares Placenta 8to9weeks 2NHP8to9W"
 /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5',
 TGTTACCAACTGAAGTGGGAGCGCGGATTTTTTTTTTTT 3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT7T3 vector
 (Pharmacia). Library constructed by Bento Soares and
 M.Fatima Bonaldo."

FEATURES

source
 1..430
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:3887512"
 /db_xref="taxon:9606"
 /clone="IMAGE:257902"
 /dev_stage="two placentae: one from 8 weeks and another
 from 9 weeks post conception"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares Placenta 8to9weeks 2NHP8to9W"
 /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5',
 TGTTACCAACTGAAGTGGGAGCGCGGATTTTTTTTTTTT 3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT7T3 vector
 (Pharmacia). Library constructed by Bento Soares and
 M.Fatima Bonaldo."

BASE COUNT 88 a 115 c 85 g 141 t 1 others

ORIGIN

Query Match 39.8%; Score 266.4; DB 14; Length 430;
 Best Local Similarity 81.5%; Pred. No. 4.4e-59;
 Matches 353; Conservative 0; Mismatches 57; Indels 23; Gaps 3;
 QY 225 TGAGTTGGCGAGAGATCAACATATAGGCTTAGCCGAGAGAGAAAGTTTACAGCTCTCTG 284
 Db 430 TGATTTGGCGAGAGATCAACATATAGGCTTAGCCGAGAGAGAGTTTACACCTCTCTG 371
 QY 285 AGCTGATTTGGGCTATGCTTACTGGCTCCCTTTGTCCAGGAAACCCACTGATGAGAGC 344
 Db 370 AGCTGATTTGGGCTATGCTT-----GGAGCCCGGTGATGAGGAGC 332
 QY 345 CTAAGAGAGAGAAACCCACC-CACTAAAGTCCGAATCTCTACACCTGATCAGAGAGAGAA 403
 Db 331 CTCAGCAAGNGAACCCACCAACTGAAAGTCCGGATCTTGACCTGGTCAGAGAGAGAA 272
 QY 404 GATGATCAGGGTGCAGCTGAGATTCAAGTGCCTGACCTGGAAAGCCGATCTCCAGAGACTA 463
 Db 271 GAAGATCAGGGTGCAGCTGAGATCAAGTGCCTGACCTGGAAAGCTGATCTCCAGGAGCTG 212

```
QY 464 TGTCAACAAGACTGGGATGGATGTGAAGTGGTACTGTATGTCAACGGGAAGATTCTA 523
Db |||||
QY 211 TCTCAGTCAAAAGACTGGGGTGAATGTGAAATGGTCTCTGATGACCCAGGGGAAGATTCTG 152
Db |||||
QY 524 CCAAAAGCAGAGCCTTTTAAATGCCAGAAGCAGGTGAAGGGAAATCACACGGTTTAAAGG 583
Db |||||
QY 151 CCAAAATCAGAACAAATTTTAAATGCCAGAAGAGGTGACAGGCAACCAAGGTTTAAATG 92
Db |||||
QY 584 AAGATAAGCTGAACAACAC-AAACTGTTTTTATATTAGATATTTTACTTTAAAGAGTCT 642
Db |||||
QY 91 AAGACAAGCTGAACAACACAAAACCTGTTTTTATCTAAGATATTGACTTAAATAATCG 32
Db |||||
QY 643 TAATTAATTTTG 655
Db |||||
QY 31 AAATAAATTTTG 19
Db |||||
```

Search completed: January 18, 2004, 18:36:19
Job time : 1023.55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: April 25, 2004, 07:00:48 ; Search time 4256.62 Seconds
(without alignments)
16458.288 Million cell updates/sec

Title: US-10-051-835-25
Perfect score: 2346
Sequence: 1 gccagaggggaaaaaagag.....aaacattgatgattacat 2346

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:

- 1: em_estba:*
- 2: em_esthm:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hcc:*
- 9: gb_estl:*
- 10: gb_est2:*
- 11: gb_hcc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estun:*
- 16: em_eston:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_phg:*
- 27: em_gss_vrl:*
- 28: gb_gssl:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	947.8	40.4	1082	12	BM547723
C	742.4	31.6	762	12	BM977480
C	719.4	30.7	799	12	BM857800
4	692.4	29.5	806	12	BM291471

5	689	29.4	842	12	BI869486
6	688	29.3	704	12	BQ014898
7	687.6	29.1	764	12	BG177619
8	677.6	28.9	712	10	AW961203
9	665.8	28.4	737	14	CD521204
10	658	28.0	729	14	CF145565
11	642.4	27.4	644	14	CB132354
12	632	26.9	681	14	CF146273
13	625	26.6	948	10	BF796780
14	619	26.4	647	12	BM668540
15	610.8	26.0	953	12	EG254117
16	603.8	25.7	624	10	AW388283
17	597	25.4	597	14	CB138744
18	594	25.3	885	10	BE545606
19	593.6	25.3	600	10	AW388333
20	589	25.1	679	10	AW388455
21	586.2	25.0	643	12	BG944108
22	583.4	24.9	585	14	CB116564
23	580.4	24.7	608	10	AW388339
24	578.6	24.7	710	14	CD640940
25	576.8	24.6	590	12	BM740046
26	573.6	24.5	589	10	AW388414
27	573.2	24.4	631	13	BQ807293
28	573	24.4	573	14	CD677372
29	573	24.4	629	10	AW388273
30	569.2	24.3	660	10	AW378440
31	562.4	24.0	613	14	CF146275
32	562	24.0	578	13	BQ678642
33	560.2	23.9	605	10	AW388336
34	551.6	23.5	567	10	AW388480
35	544.4	23.2	588	13	BQ364149
36	544	23.2	595	10	AW388453
37	543.6	23.2	597	10	AW388607
38	540.2	23.0	608	10	AW388413
39	538.4	22.9	589	14	CF145360
40	537.2	22.9	611	10	AW388642
41	534.8	22.8	673	13	BQ364178
42	530.6	22.6	565	10	AW388711
43	529	22.5	567	10	AW388561
44	528.8	22.5	532	14	CB135713
45	526	22.4	526	10	AW237145

ALIGNMENTS

RESULT 1
BM547723
LOCUS
DEFINITION
AGENCOURT_6507143 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5727848
5', mRNA sequence.
ACCESSION
BM547723
VERSION
BM547723.1 GI:18781736
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1082)
NIH-MGC http://imgc.ncbi.nlm.nih.gov/.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Invitrogen
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12722 row: j column: 09
High quality sequence start: 3

High quality sequence stop: 723.

FEATURES

Location/Qualifiers
1..1082
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5727848"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH MGC 124"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: EcoRV (destroyed); Site 2: NotI; RNA source male hippocampus, age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 0.9-4 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 012."

ORIGIN

Query Match 40.4%; Score 947.8; DB 12; Length 1082;

Best Local Similarity 96.4%; Pred. No. 3.5e-180; Indels 7; Gaps 6;

Matches 1031; Conservative 0; Mismatches 31; Indels 7; Gaps 6;

378 GAACAGATTGTTCCCATGAATGTATGCTCATCTAGCTAGAAACGATCTATGTTAAAT 437
16 GATACAGATTGTTCCCATGAATGTATGCTCATCTAGCTAGAAACGATCTATGTTAAAT 75
438 GACTGTATATGAAATATTTCAAGTACTACCCCAATCACTTCTTATGCTCTGAAAG 497
76 GACTGTATATGAAATATTTCAAGTACTACCCCAATCACTTCTTATGCTCTGAAAG 135
498 AAGAAAGCAATGAAATCACTATGATTTATGCAACAAACAGAAATCTCCAAAT 557
136 AAGAAAGCAATGAAATCACTATGATTTATGCAACAAACAGAAATCTCCAAAT 195
558 TTAAGTAATCTGCTCTCTCTGAGAAATGTTACCTATAGTATTTTCTCTATGAT 617
196 TTAAGTAATCTGCTCTCTCTGAGAAATGTTACCTATAGTATTTTCTCTATGAT 255
618 GTTATTACTACTGCTTAATCAATCAATTTCTATAAATTTCTACTTAAGTCCTTAAGAACTG 677
256 GTTATTACTACTGCTTAATCAATCAATTTCTATAAATTTCTACTTAAGTCCTTAAGAACTG 315
678 GGTCTCTCTCTGATGTTATGATGTTATGTTAGAAAGAAACAACTTTACTCTTTAGGAC 737
316 GGTCTCTCTCTGATGTTATGATGTTATGTTAGAAAGAAACAACTTTACTCTTTAGGAC 375
738 AATTCTCTAGAACTCTATAGTATGATCAGGATATATTTTGTCTTAAATATATTTTGTAT 797
376 AATTCTCTAGAACTCTATAGTATGATCAGGATATATTTTGTCTTAAATATATTTTGTAT 435
798 TTGGAATACAGACTTGGCTCCAAATTTCACTCTTTGCAATAGTATGACTTTCACTA 857
436 TTGGAATACAGACTTGGCTCCAAATTTCACTCTTTGCAATAGTATGACTTTCACTA 495
858 GAATCTCTCAACATTTGGCAACTTTGCAATATGAGCATCATATGTTGTAAGCTCTATC 917
496 GAATCTCTCAACATTTGGCAACTTTGCAATATGAGCATCATATGTTGTAAGCTCTATC 555
918 ATTATAATGCTATGATGATACATTTGTTTCTCCCTATGCAACAGGTGAACAACTAGTT 977
556 ATTATAATGCTATGATGATACATTTGTTTCTCCCTATGCAACAGGTGAACAACTAGTT 615
978 GTTTTTTACTGATCTAAATGTTGGCTACTGTTGATTTTATAGTATGCACTGCAAAA 1037
616 GTTTTTTACTGATCTAAATGTTGGCTACTGTTGATTTTATAGTATGCACTGCAAAA 675
1038 AAGCAAGACAAATGGCTCTTGTACTGAAATACCTTGGCAAACTTATTTGGCTCTTCATTT 1097
676 AAGCAAGACAAATGGCTCTTGTACTGAAATACCTTGGCAAACTTATTTGGCTCTTCATTT 735
1098 TCTGACAGACAGGATTTGACTCAATA-TTTGTAGAGCTTCGGTAGAATGATACATGTT 1156

Db 736 TCTGACAGACAGATTTGACTCAATATTTGTAGAGCTTCGCTAGATGATACATGTT 795
Qy 1157 AGTGATGCACCTGGTAGAAATGTTTTAGTTATGACTCAGAAATCACTCAGGATGAT 1216
Db 796 AGTGATGCACCTGGTAGAAATGTTTTAGTTATGACTCAGAAATCACTCAGGATGAAA 855
Qy 1217 CTTTTATGCTTTTTTATTTGTAAGCATATCTGAATTTACTTTTATAAGATGGTTTTAGAAA 1276
Db 856 CTTTTATGCTTTTTTATTTGTAAGCATATCTGAATTTACTTTTATAAGATGGTTTTAGAAA 914
Qy 1277 GCTTTCTCTAAAAATTTGGCTAGGAATGGTAACCTTCATTTTCAGTTCGCAAGGGGTAGA 1336
Db 915 GCTTTCTCTAAAAA-TTGGCTAGGAATGGTAACCTTCATTTTCANNTGCCAGGGGGTAA 973
Qy 1337 AAAATATATGCTGTTGTTAT-GTTTATGTTAAACATATATTA-GGTACTATCTATGAA 1394
Db 974 AAAATATATGCGGGGTATGTTAGGTTAACTTATTAAGGTTACTATTAATGATGAA 1033
Qy 1395 TGTATTTAAATA--TTTTTCATATCTCTGACAGCAATTTTAAATTTGC 1441
Db 1034 AGTAATTAATAATTTTCTAAATCTGGGAAACCAATTTATAATTTGC 1082

RESULT 2

BM977480/c
LOCUS 762 bp mRNA linear EST 21-FEB-2003
DEFINITION UI-CF-EN1-ase-b-22-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
UI-CF-EN1-ase-b-22-0-UI 3', mRNA sequence.

ACCESSION

BM977480

VERSION

BM977480.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 762)

AUTHORS

Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE

Normalization and subtraction: two approaches to facilitate gene

JOURNAL

Genome Res. 6 (9), 791-806 (1996)

MEDLINE

97044477

PUBMED

8889548

COMMENT

McCrack Lab

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com) or from Open Biosystems

(www.openbiosystems.com).

The following repetitive elements were found in this cDNA

sequence: 1-64, >AT rich#Low_complexity

Seq primer: M13 FORWARD

POLYA=Yes.

FEATURES

Location/Qualifiers

1..762

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clones="UI-CF-EN1-ase-b-22-0-UI"

/tissue_type="Primary Lung Cystic Fibrosis Epithelial

Cells"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-CF-EN1"

/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-EN1 is a normalized cDNA library containing the following tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CTGCTCAGGT. TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS 6hr to LPS 24h TAG LIB=UI-CF-EN1 TAG_SEQ=CTGCTCAGGT"

ORIGIN

Query Match 31.6%; Score 742.4; DB 12; Length 762;
Best Local Similarity 99.9%; Pred. No. 6.5e-139;
Matches 743; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1603 CATTCGCCAGATTTCAGGGGATCTGGGTTTGGTCTAGTGAATGAAACACAAAGACAG 1662
DB 762 CATTCGCCAGATTTCAGGGGATCTGGGTTTGGTCTAGTGAATGAAACACAAAGACAG 703
QY 1663 AGAGATCCAGCTGAAAAGAGTGAATCCTCAATATCCTAACTAACTGTCCTCAACTCAAG 1722
DB 702 AGAGATCCAGCTGAAAAGAGTGAATCCTCAATATCCTAACTAACTGTCCTCAACTCAAG 643
QY 1723 CAGAGTTTCTTCACCTCTGCACTGTGATCATGAACTTAGTAGAGGGATGTGTGTAAT 1782
DB 642 CAGAGTTTCTTCACCTCTGCACTGTGATCATGAACTTAGTAGAGGGATGTGTGTAAT 583
QY 1783 TTATACAAATTAATACAAATGCTTCAATGATGATAAAATCTTAAGACAAACTGCAAT 1842
DB 582 TTATACAAATTAATACAAATGCTTCAATGATGATAAAATCTTAAGACAAACTGCAAT 523
QY 1843 TTATTTCTGATCCATCCATTCATATGAACTAAGATATTTATCTATGAAGATATA 1902
DB 522 TTATTTCTGATCCATTCATATGAACTAAGATATTTATCTATGAAGATATA 463
QY 1903 AATGTTGACAGAGACTTTTCATCTGTGGATTCGGTGTCTTAGGGTTCCTAGCACTGA 1962
DB 462 AATGTTGACAGAGACTTTTCATCTGTGGATTCGGTGTCTTAGGGTTCCTAGCACTGA 403
QY 1963 TGCCCTGCAAGCATGTGATATGTGAATAAATGATTTCTTCTATAGCTAATCAGTTC 2022
DB 402 TGCCCTGCAAGCATGTGATATGTGAATAAATGATTTCTTCTATAGCTAATCAGTTC 343
QY 2023 CCTCTGGGGAGAGTTCCTGCTACTGCAATCACAATGCCAGATGGTGTATATGGGCTATTG 2082
DB 342 CCTCTGGGGAGAGTTCCTGCTACTGCAATCACAATGCCAGATGGTGTATATGGGCTATTG 283
QY 2083 TGTAGTAAAGTGAAGTCTATGAGTAAGTGTGTGTCTTCTATAGCTAATCAGTTC 2142
DB 282 TGTAGTAAAGTGAAGTCTATGAGTAAGTGTGTGTCTTCTATAGCTAATCAGTTC 223
QY 2143 TTGATGCAATGCTTTTGTATGGAATAAATTTTGGTGCAATATGATGTCATTCATCTTTG 2202
DB 222 TTGATGCAATGCTTTTGTATGGAATAAATTTTGGTGCAATATGATGTCATTCATCTTTG 163
QY 2203 CATTGAATGAATTTGGTGTATTTATGATATATATCTGCTCAGGCTCTAGTGTCTT 2262
DB 162 CATTGAATGAATTTGGTGTATTTATGATATATATCTGCTCAGGCTCTAGTGTCTT 103
QY 2263 CAACCAATTTATACCAATTTTGTACATATTTTACCTTGAATAATTTTAAATGGAATTT 2322
DB 102 CAACCAATTTTATACCAATTTTGTACATATTTTACCTTGAATAATTTTAAATGGAATTT 43
QY 2323 AATAAACAATTTGATGTTTACAT 2346

Db 42 AATAAACAATTTGTTAGTTTACAT 19

RESULT 3
BI857800/c
LOCUS BI857800.1
DEFINITION Homo sapiens cDNA clone IMAGE:5397469 5', mRNA linear EST 10-OCT-2001
ACCESSION BI857800
VERSION BI857800.1
KEYWORDS GI:15998547
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 799)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLNL2013 row: d column: 14
High quality sequence stop: 776.
Location/Qualifiers
FEATURES
source
1..799
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5397469"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="NIH MGC 87"
/clone_lib="NIH MGC 87"
/notes="Organ: breast; Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
ORIGIN
Query Match 30.7%; Score 719.4; DB 12; Length 799;
Best Local Similarity 97.4%; Pred. No. 2.7e-134;
Matches 753; Conservative 0; Mismatches 16; Indels 4; Gaps 2;

QY 1571 GGATCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1630
DB 799 GGATCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 740
QY 1631 GTTTGGTCTAGTGAATGAAACACAAGAACAGAGATGTC---AGCTGAAAAAGATGAT 1687
DB 739 GTTTGGTCTAGTGAATGAAACACAAGAACAGAGATGTC---AGCTGAAAAAGATGAT 680
QY 1688 CCTCAATATCCTTAACCTAAGTGTCTCACTCAAGCAGAGTCTTCTCACTCTGCACTGT 1747
DB 679 CCTCAATATCCTTAACCTAAGTGTCTCACTCAAGCAGAGTCTTCTCACTCTGCACTGT 620
QY 1748 GATCATGAAACTTAGTAGAGGGGATGTTGTGTATTTTATACAAATTTAATACATGTCCT 1807
DB 619 GATCATGAAACTTAGTAGAGGGGATGTTGTGTATTTTATACAAATTTAATACATGTCCT 560
QY 1808 ACATTGATAAATTTTAAAGAGCAAACTGCAATTTTA-TTTCTGATCCACATTCCTCAAT 1866
DB 559 ACATTGATAAATTTTAAAGAGCAAACTGCAATTTTGAAGTTTCTGCAATCCATTCCTCA 500
QY 1867 CATATTGAACTAAGATATTTATCTATGAAGATATAAATGTTGTCAGAGAGACTTTTCTCT 1926
DB 499 CATATTGAACTAAGATATTTATCTATGAAGATATAAATGTTGTCAGAGAGACTTTTCTCT 440

```

QY 1927 GTGATTCGGTGTCTTCTAGGGTTCTAGCACTGATGCTGCACAGCAATGTCATATGT 1986
D 439 GTGATTCGGTGTCTTCTAGGGTTCTAGCACTGATGCTGCACAGCAATGTCATATGT 380
QY 1987 GAAATATAAATCGATTCCTATACCTAAATCAGTTCCTCTGGGAGAGTTCCTGGTACTG 2046
D 379 GAAATATAAATCGATTCCTATACCTAAATCAGTTCCTCTGGGAGAGTTCCTGGTACTG 320
QY 2047 CAATCAAAATGCCAGATGGTGTATTTATGGCTATTTGTGTAAGTAAGTGAAGATGCTAT 2106
D 319 CAATCAAAATGCCAGATGGTGTATTTATGGCTATTTGTGTAAGTGAAGTGAAGATGCTAT 260
QY 2107 GAAGTAAGTGTGTGTCTTATCATCTTATGAACTTCTGATGCTGCTTGTGTAAGTGA 2166
D 259 GAAGTAAGTGTGTGTCTTATCATCTTATGAACTTCTGATGCTGCTTGTGTAAGTGA 200
QY 2167 ATAAATTTGGTGCATATGATGTCATTCATTCATTTGCAATTTGAATTTGGTGTAT 2226
D 199 ATAAATTTGGTGCATATGATGTCATTCATTCATTTGCAATTTGAATTTGGTGTAT 140
QY 2227 TTATATGATATATACCTGTCACGCTTCTAGTTCCTCAACATTTTATAACATTTTGT 2286
D 139 TTATATGATATATACCTGTCACGCTTCTAGTTCCTCAACATTTTATAACATTTTGT 80
QY 2287 ACATATTTTACTTGAAATATTTTAAATGGAATTTTAAATGGAATTTTAAATGGAATTTTGTAG 2339
D 79 ACATATTTTACTTGAAATATTTTAAATGGAATTTTAAATGGAATTTTAAATGGAATTTTGTAG 27

RESULT 4
LOCUS BG291471
DEFINITION 602387358F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4516167 5',
mRNA sequence.
ACCESSION BG291471
VERSION BG291471.1 GI:13049386
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgs@nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M10407 row: C column: 16
High quality sequence stop: 724.
FEATURES
1..806
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4516167"
/tissue_type="transitional cell papilloma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_93"
/notes="Organ: bladder; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
ORIGIN

```

```

Query Match 29.5%; Score 692.4; DB 12; Length 806;
Best Local Similarity 96.6%; Pred. No. 7.1e-129;
Matches 761; Conservative 0; Mismatches 21; Indels 6; Gaps 5;

QY 1003 CTACCTGTGATTTTATAGTATGCACATGTTCAGAAAAGGCAAGACAAATGGCCCTTTGTA 1062
D 1 CTACCTGTGATTTTATAGTATGCACATGTTCAGAAAAGGCAAGACAAATGGCCCTTTGTA 60
QY 1063 CTGAATACCTTCGGCAAACTTATTTGGGCTTCATTTCTGACAGACAGGATTTGACTCAAT 1122
D 61 CTGAATACCTTCGGCAAACTTATTTGGGCTTCATTTCTGACAGACAGGATTTGACTCAAT 120
QY 1123 ATTGTAGAGCTTCGGTGAAGTGAATACATGCTAGTGAATGACCTGGTAGAAATGGTTTT 1182
D 121 ATTGTAGAGCTTCGGTGAAGTGAATACATGCTAGTGAATGACCTGGTAGAAATGGTTTT 180
QY 1183 TAGTTATTTGACTCAGAAATTCATCTCAGGATGAATCTTTTATGCTTTTATTTGTAAGCAT 1242
D 181 TAGTTATTTGACTCAGAAATTCATCTCAGGATGAATCTTTTATGCTTTTATTTGTAAGCAT 240
QY 1243 ATCTGAATTTACTTTATTAAGATGCTTTTAGAAAAGCTTTGCTAAAAATTTGGCCTAGGA 1302
D 241 ATCTGAATTTACTTTATTAAGATGCTTTTAGAAAAGCTTTGCTAAAAATTTGGCCTAGGA 300
QY 1303 ATGCTAACTTCATTTTCAAGTTCGCAAGGGGTAGAAAATTAATATGCTGTGTTTATGTTT 1362
D 301 ATGCTAACTTCATTTTCAAGTTCGCAAGGGGTAGAAAATTAATATGCTGTGTTTATGTTT 360
QY 1363 ATGTTAACAATTTATTTAGTACTATCTATGAATGTTTAAATTTTCAATTTCTGTG 1422
D 361 ATGTTAACAATTTATTTAGTACTATCTATGAATGTTTAAATTTTCAATTTCTGTG 420
QY 1423 ACAAGCATTTTATTTTGAACAGTGGAGTCCCATTTAGCCAGTGGGAAAGTCTTCGAA 1482
D 421 ACAAGCATTTTATTTTGAACAGTGGAGTCCCATTTAGCCAGTGGGAAAGTCTTCGAA 480
QY 1483 CTCAGGTTACCCCTTGAAGGATATGCTGGCAGCCATCTCTTTGATCTGTGTTAAACTGTA 1542
D 481 CTCAGGTTACCCCTTGAAGGATATGCTGGCAGCCATCTCTTTGATCTGTGTTAAACTGTA 540
QY 1543 ATTATAGACAGCTTAATCCCTAACTTGGATCTGGAATCATTAGTATGACCTTCGTAC 1602
D 541 ATTATAGACAGCTTAATCCCTAACTTGGATCTGGAATCATTAGTATGACCTTCGTAC 600
QY 1603 CATTCCCAAGATTTACAGGGGCATCGTGGGCTTGGTCTAAAGTGAATGAAACCCCAAGAA 1659
D 601 CATTCCCAAGATTTACAGGGGCATCGTGGGCTTGGTCTAAAGTGAATGAAACCCCAAGAA 660
QY 1660 CAGAGATCCAGCTGAAAAGAGTGAATCTCAATATCCCTAACTTAACCTGCTCAACTC 1719
D 661 CAGAGATCCAGCTGAAAAGAGTGAATCTCAATATCCCTAACTTAACCTGCTCAACTC 719
QY 1720 AAGCAGAGTTTCTTCACTCTGGCAGCTGATCATGAACTTAGTAGAGGGGATTTGTGT 1779
D 720 AAGCAGAGTTTCTTCACTCTGGCAGCTGATCATGAACTTAGTAGAGGGGATTTGTGT 777
QY 1780 ATTATTA 1787
D 778 CTTTCAAA 785

RESULT 5
LOCUS BI869486
DEFINITION 603392919F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5402885 5',
mRNA sequence.
ACCESSION BI869486
VERSION BI869486.1 GI:16043159
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```



```

Db      600  CGTGTTCCTTAGGCTCCTAGCACTGATGCTGTCACAGCATGTGATATGTGAATAACA 659
Qy      1995  ATGATTCCTTATAGCTAAATGAGTTCCTCTGCGGAGAGATTCCTGTAATGCAATCACA 2054
Db      660  ATGATTCCTTATAGCTAAATGAGTTCCTCTGCGGAGAGATTCCTGTAATGCAATCACA 719
Qy      2055  ATGCCAGATGGTGTATATGGGCTAT 2079
Db      720  ATGCCAATGGTGTATGGGCAAT 744

RESULT 8
LOCUS   AW961203                712 bp    mRNA    linear    EST 01-JUN-2000
DEFINITION EST373275 MAGE resequences, MAGEF Homo sapiens cDNA, mRNA sequence.
ACCESSION AW961203
VERSION   AW961203.1 GI:8150887
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 712)
AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
          Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
          Quackenbush,J.
          Assessment of gene expression patterns in a model of colon tumor
          metastasis using a 19,200 element cDNA microarray
          Unpublished (2000)
          Contact: John Quackenbush
          The Institute for Genomic Research
          9712 Medical Center Dr., Rockville, MD 20850, USA
          Tel: 301 838 3528
          Fax: 301 838 0208
          Email: johnq@tigr.org
          Plate: 155
          Seq primer: Reverse.
          Location/Qualifiers
FEATURES             source
     1..712
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone_lib="MAGE resequences, MAGEF"
     /note="Vector: pBluescriptSKm"

ORIGIN
Query Match      28.9%; Score 677.6; DB 10; Length 712;
Best Local Similarity 99.2%; Pred. No. 6.7e-126;
Matches 702; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Qy      29  AGGTATCTCTTTGCAAGTGGTCACTGTATTTTGGTACCTTCTGTCAGAGGTATTATT 88
Db      1  AGGTATCTCTTTGCAAGTGGTCACTGTATTTTGGTACCTTCTGTCAGAGGTATTATT 59
Qy      89  ACAGCATCTTGTGGGAAACCTATTAGGCTTTGATGCTTTAAAGTGTATATTGTTGG 148
Db      60  ACAGCATCTTGTGGGAAACCTATTAGGCTTTGATGCTTTAAAGTGTATATTGTTGG 119
Qy      149  GTTGTGAGTGGTCTGACTTAATGTGTATTAATAAATTTAGACATCAAAATTTCTCTACTA 208
Db      120  GTTGTGAGTGGTCTGACTTAATGTGTATTAATAAATTTAGACATCAAAATTTCTCTACTA 179
Qy      209  ACTAATCTTTTAGATGCATCTTGAAGCAGACAGTCATATCATCTGGGAGCAATGCAA 268
Db      180  ACTAATCTTTTAGATGCATCTTGAAGCAGACAGTCATATCATCTGGGAGCAATGCAA 239
Qy      269  TGTGGTTACTGCTCTAGTGTGAGTCTCTTATTTCAAGAGATTTCTGAATTAATTT 328
Db      240  TGTGGTTACTGCTCTAGTGTGAGTCTCTTATTTCAAGAGATTTCTGAATTAATTT 299
Qy      329  TCCTTAGAATTTCTCTCTTCAATTCACCAAGTACAAACATCTTTGAAGAATGAACAGATTG 388
Db      300  TCCTTAGAATTTCTCTCTTCAATTCACCAAGTACAAACATCTTTGAAGAATGAACAGATTG 359

```

```

Qy      389  TTCCCATGAATGATGCTCATACTCGACTAGAAAAGATCTATGTTAAATGACTGTGTATA 448
Db      360  TTCCCATGAATGATGCTCATACTCGACTAGAAAAGATCTATGTTAAATGACTGTGTATA 419
Qy      449  TGAATTTATTTCAAGTACTACCCCAATAACTTTCTTTATTTGCTCTGAAAGAGAAAGCAA 508
Db      420  TGAATTTATTTCAAGTACTACCCCAATAACTTTCTTTATTTGCTCTGAAAGAGAAAGCAA 479
Qy      509  TGTAAATCACTATGATTATTGACAAACCAACAGAAATTCCTCAACAATTTTAAAGTAATCT 568
Db      480  TGTAAATCACTATGATTATTGACAAACCAACAGAAATTCCTCAACAATTTTAAAGTAATCT 539
Qy      569  GATCCCTCTTTCTGGAGAAAATTTGTTACCTTAATAGTATTTTCTTATGAATGTTATTACTAC 628
Db      540  GATCCCTCTTTCTGGAGAAAATTTGTTACCTTAATAGTATTTTCTTATGAATGTTATTACTAC 599
Qy      629  TGGTATAAATCAAAATTTCTATAAATTTCTTAAATTTCTTAAGTCTTAAGAACTGGTTCCTTT 688
Db      600  TGGTATAAATCAAAATTTCTATAAATTTCTTAAATTTCTTAAATTTCTTAAACTGGTTCCTTT 659
Qy      689  TGATGTTATTCATGTTTCAGAAA-GGAAACAACACTTTACTCTTTTAGG 735
Db      660  TGATGTTATTCATGTTTCAGAAAGGAGAAACACACTTTACTTTTAGG 707

RESULT 9
LOCUS   CD521204                737 bp    mRNA    linear    EST 06-JUN-2003
DEFINITION AGENCOURT 14353302 NIH MGC 191 Homo sapiens cDNA clone
IMAGE: 30412131 5', mRNA sequence.
ACCESSION CD521204
VERSION   CD521204.1 GI:31452922
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 737)
AUTHORS   NIH-MGC http://mgs.nci.nih.gov/
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Daniela S. Gerhard, Ph.D.
          Office of Cancer Genomics
          National Cancer Institute / NIH
          Bldg. 31 Rm10A07 Bethesda, MD 20892
          Email: sgabs-x@mail.nih.gov
          Tissue Procurement: Narayan Bhat
          cDNA Library Preparation: CLONTECH Laboratories, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: NDCM205 row: a column: 04
          High quality sequence stop: 592.
          Location/Qualifiers
FEATURES             source
     1..737
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone_lib="IMAGE:30412131"
     /tissue_type="Pooled"
     /lab_host="DH10B (T1 phage-resistant)"
     /clone_lib="NIH MGC 191"
     /note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);
     Site 2: SfiI (ggcgccctggcc); Library is oligo-dT primed
     and directionally cloned. pBMC - peripheral Blood
     Mononuclear Cells. RNA was pooled from 3/6hour stimulation
     with PMA adn Ionomycin. 5' and 3' adaptors were used in
     cloning as follows: 5' adaptor sequence:
     5'-CACGGCATTATGCC-3' and 3' adaptor sequence:
     5'-ATTCTAGGCGCGAGCGCCGACATG-dT(30)BN-3' (where B = A,

```

C, or G and N = A, C, G, or T). Average insert size 1.69 kb (range 0.70-5.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN		Query Match		28.4%; Score 665.8; DB 14; Length 737;	
		Best Local Similarity		99.7%; Pred. No. 1.6e-123;	
		Matches 667; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
QY	1033	AGAAAAAGGCAAGCAAAATGGCTCTGTACTGAATACCTTCGGCAAACTTATTGGCTCTT	1092		
Db	3	AGAAAAAGGCAAGCAAAATGGCTCTGTACTGAATACCTTCGGCAAACTTATTGGCTCTT	62		
QY	1093	CATTTCCTGACAGACAGATTTGACTCAATATTGTAGACCTTCGGTAGAATGATACA	1152		
Db	63	CATTTCCTGACAGACAGATTTGACTCAATATTGTAGACCTTCGGTAGAATGATACA	122		
QY	1153	TGGTAGTGATGACCTGGTAGAAATGGTTTTAGTTATTGACTCAGAAATTCATCTCAGGAT	1212		
Db	123	TGGTAGTGATGACCTGGTAGAAATGGTTTTAGTTATTGACTCAGAAATTCATCTCAGGAT	182		
QY	1213	GAATCTTTTATGCTCTTTTATTTGTAAGCATATCTGAATTTACTTTATAAGATGGTTTA	1272		
Db	183	GAATCTTTTATGCTCTTTTATTTGTAAGCATATCTGAATTTACTTTATAAGATGGTTTA	242		
QY	1273	GAAAGCTTTGTCTAAAAATTTGGCTAGGAATGTAATCTTCATTTTCAGTTGGCAAGGGG	1332		
Db	243	GAAAGCTTTGTCTAAAAATTTGGCTAGGAATGTAATCTTCATTTTCAGTTGGCAAGGGG	302		
QY	1333	TAGAAAAATAATATGTGTGTGTATGTTTATGTTATGTTAAATATTTAGGTACTATCATG	1392		
Db	303	TAGAAAAATAATATGTGTGTGTATGTTTATGTTTATGTTTAAATATTTAGGTACTATCATG	362		
QY	1393	AATGATTTAATAATTTTTCATATCTGTGACAGCATTTATATTTTGCAACAGTGGAG	1452		
Db	363	AATGATTTAATAATTTTTCATATCTGTGACAGCATTTATATTTTGCAACAGTGGAG	422		
QY	1453	TCCATTTAGCCAGTGGGAAAGTCTTGAACTCAGTTACCTTTGAAAGGATATGCTGGCA	1512		
Db	423	TCCATTTAGCCAGTGGGAAAGTCTTGAACTCAGTTACCTTTGAAAGGATATGCTGGCA	482		
QY	1513	CCCATCTTTTATGCTGTCTTAACTGTAATTTATAGACAGCTAAATCCCTAACTTGG	1572		
Db	483	CCCATCTTTTATGCTGTCTTAACTGTAATTTATAGACAGCTAAATCCCTAACTTGG	542		
QY	1573	ATCTGGAATGATTAGTTATGACCTTTGACCTTCCATTTCCAGAAATTTCCAGGGGATGTTGGT	1632		
Db	543	ATCTGGAATGATTAGTTATGACCTTTGACCTTCCATTTCCAGAAATTTCCAGGGGATGTTGGT	602		
QY	1633	TGGTCTAGTATTGAACACACAGACAGAGAGATCCAGCTGAAAGAGATGATCCTCA	1692		
Db	603	TGGTCTAGTATTGAACACACAGACAGAGATCCAGCTGAAAGAGATGATCCTCA	662		
QY	1693	ATATCCTAA 1701			
Db	663	ATATCCTAA 671			

RESULT 10
CF145565/c
LOCUS
DEFINITION
UI-HF-CB0-asq-e-04-0-UI.r1 NIH_MGC_210 Homo sapiens CDNA clone
IMAGE:30569331 5', mRNA sequence.
ACCESSION
CF145565
VERSION
CF145565.1 GI:33261009
EST.
KEYWORDS
Homo sapiens (human)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 729)

AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Tim Ratliff
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/humanfl.html
The following repetitive elements were found in this cDNA
sequence: 27-49, >POLY_A#simple_repeat (matched complement)
Seq primer: PYX-5.

FEATURES
source

Location/Qualifiers
1..729
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30569331"
/tissue_type="CNCAP(3)T-225 cell line"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_MGC_210"
/note="Organ: Prostate; Vector: p7T3 Pac; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with Not I and then cloned
directionally into p7T3 Pac vector. The library tag
sequence located between the Not I site and the polyA tail
is CCCAC. Tissue was provided by Tim Ratliff."

ORIGIN

Query Match		28.0%; Score 658; DB 14; Length 729;	
Best Local Similarity		99.7%; Pred. No. 5.8e-122;	
Matches 680; Conservative 0; Mismatches 0; Indels 2; Gaps 2;			
QY	1665	AGATCCAGCTGAAAAGAGTGAATCCTCAATATCCTAACTAACTGGTCTCTCACTCAAGCA	1724
Db	729	AGATCCAGCTGAAAAGAGT-ATCCTCAATATCCTAACTAACT-GTCTCACTCAAGCA	672
QY	1725	GAGTTTCTTCTACTCTGGCACTGTGATCATGAACTTAGTAGAGGGGATGTGTATTTT	1784
Db	671	GAGTTTCTTCTACTCTGGCACTGTGATCATGAACTTAGTAGAGGGGATGTGTATTTT	612
QY	1785	ATACAAATTAATAACAATGTCTTACATTTGATTAATTTCTTAAGAGCAAACTGATTTT	1844
Db	611	ATACAAATTAATAACAATGTCTTACATTTGATTAATTTCTTAAGAGCAAACTGATTTT	552
QY	1845	ATTTCCTGCATCCACATTCCTCAATCATATTAGAACTATTTATCTATGAAGATATAA	1904
Db	551	ATTTCCTGCATCCACATTCCTCAATCATATTAGAACTATTTATCTATGAAGATATAA	492
QY	1905	TGGTGACAGAGACATTTTCACTGTGGATTTGCTTTTCTTAGGGTTCTTAGCACTGATG	1964
Db	491	TGGTGACAGAGACATTTTCACTGTGGATTTGCTTTTCTTAGGGTTCTTAGCACTGATG	432
QY	1965	CCTGACACAGCATGTGATATGTAATAAATGGAATCTTCTATAGCTAAATAGTTCCTC	2024
Db	431	CCTGACACAGCATGTGATATGTAATAAATGGAATCTTCTATAGCTAAATAGTTCCTC	372
QY	2025	TCTGGGGAGAGTCTTGGTACTGCTCAATCAATGCCAGATGGTGTATTATGGGCTATTGTG	2084

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/humanfl.html>
The following repetitive elements were found in this cDNA
sequence: 27-49, >Pou1f_AHSimple_repeat (matched complement)
Seq primer: pYX-5.

FEATURES

Location/Qualifiers
1..681
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30569820"
/tissue_type="CNAP(3)T-225 cell line"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH MGC 210"
/notes="Organ: Prostate; Vector: pTV73 Pac; Site 1: Ecor I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with Ecor I adaptor, digested with Not I and then cloned
directionally into pTV73 Pac vector. The library tag
sequence located between the Not I site and the polyA tail
is CCCAC. Tissue was provided by Tim Ratliff."

ORIGIN

Query Match 26.9%; Score 632; DB 14; Length 681;
Best Local Similarity 100.0%; Pred. No. 9.7e-117;
Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1715 AACTCAGCAGAGTTCTTCCACTGCGACATCATCAAACTAGTAGAGGGATTG 1774
Db AACTCAGCAGAGTTCTTCTGCTGCGATGATCAATGAACTAGTAGAGGGATTG 622
1775 TGTGTAATTTATACAAATTTAATAACAATGCTTACATGATAAAATCTTAAAGACAAA 1834
Db TGTGTAATTTATACAAATTTAATAACAATGCTTACATGATAAAATCTTAAAGACAAA 562
1835 ACTGCATTTATTTCTGCATCCCATTCATCATATAGACTAGACTAGATATTTATCTATG 1894
Db ACTGCATTTATTTCTGCATCCCATTCATCATATAGACTAGACTAGATATTTATCTATG 502
1895 AAGATATAAATGGTGCAGAGAGATTTTCATCTGTGGATTCGGTTTCTTACGGTTCT 1954
Db AAGATATAAATGGTGCAGAGAGATTTTCATCTGTGGATTCGGTTTCTTACGGTTCT 442
1955 AGCACTGATGCTGCAGACAGCATGTATGTGAATATAAATGATTTCTTATAGCTAA 2014
Db AGCACTGATGCTGCAGACAGCATGTATGTGAATATAAATGATTTCTTATAGCTAA 382
2015 ATGAGTTCCCTCTGGGAGAGTTCTGCTACTGCAATCACAATGCCAGATGGTTTATGG 2074
Db ATGAGTTCCCTCTGGGAGAGTTCTGCTACTGCAATCACAATGCCAGATGGTTTATGG 322
2075 GCTATTTGTGTAAGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2134
Db GCTATTTGTGTAAGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 262
2135 GGAACTCTTGATGATGATGCTTTTGTATGGAATAAATTTTGGTGCATATGATGTCAAT 2194
Db GGAACTCTTGATGATGATGCTTTTGTATGGAATAAATTTTGGTGCATATGATGTCAAT 202
2195 CAACCTTGCAATGAATGAATTTGGTTGGTATTTATATGATATATACCTGCAGCTTCT 2254
Db CAACCTTGCAATGAATGAATTTGGTTGGTATTTATATGATATATACCTGCAGCTTCT 142
2255 AGTTCCTTCAACCACTTTATACCAATTTTGTACATATTTTACTTGAATAATTTTAAAT 2314
Db AGTTCCTTCAACCACTTTTATACCAATTTTGTACATATTTTACTTGAATAATTTTAAAT 82

Qy 2315 GGAAATTTAAATAAACAATTTGATGACTTTACAT 2346
Db 81 GGAAATTTAAATAAACAATTTGATGACTTTACAT 50

RESULT 13

BF796780 948 bp mRNA linear EST 12-JAN-2001
LOCUS 602259726F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4342974 5',
DEFINITION mRNA sequence.

ACCESSION BF796780

VERSION BF796780.1 GI:12101834

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 948)

NIH-MGC <http://mhc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: ccapbs-remail.nih.gov

Tissue Procurement: Louis Staudt, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LHM9959 row: C column: 07

High quality sequence stop: 643.

Location/Qualifiers

1..948

source

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4342974"

/tissue_type="lymphoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_85"

/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 1.867 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

ORIGIN

Query Match 26.6%; Score 625; DB 10; Length 948;
Best Local Similarity 97.9%; Pred. No. 2.4e-115;
Matches 655; Conservative 0; Mismatches 10; Indels 4; Gaps 2;

Qy 14 AAAAGAGTAATGCACAGGATCTCTTTTGCAGTGGTGAATCTTATTTGAGTACCTTTGTG 73
Db 1 AAAAGAGTAATGCACAGGATCTCTTTTGCAGTGGTGAATCTTATTTGAGTACCTTTG 59
Qy 74 TGACAGGGTATTTATACAGCATCTTTGGGAAAACCTATTAGGCTTTGCAATGTTAAGC 133
Db 60 TGACAGGGTATTTATACAGCATCTTTGGGAAAACCTATTAGGCTTTGCAATGTTAAGC 119
Qy 134 TGTATAATTTGTGGTGTGAGTGGTCTGACTTAATGCTGATTAATAAATTTAGACAT 193
Db 120 TGTATAATTTGTGGTGTGAGTGGTCTGACTTAATGCTGATTAATAAATTTAGACAT 179
Qy 194 CAAATTTTCTACTAACTAACTTTATAGATGCATCTTGAAGCAGCAGTCAATCAAC 253
Db 180 CAAATTTTCTACTAACTAACTTTATAGATGCATCTTGAAGCAGCAGTCAATCAAC 239
Qy 254 TGGAGGCAATGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 313
Db 240 TGGAGGCAATGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 299
Qy 314 TTCTGAATTAATTTTCCCTAGAAATTTCTCTTCTTCCAAAGTACAAACATCACTTTGAA 373

LOCUS BG254117 953 bp mRNA linear EST 13-FEB-2001
DEFINITION 602367005F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4475233 5',
mRNA sequence.
ACCESSION BG254117
VERSION BG254117.1 GI:12763933
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 953)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-x@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10300 row: j column: 02
High quality sequence stop: 642.
FEATURES
Location/Qualifiers
1..953
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4475233"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_91"
/note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
ORIGIN
Query Match 26.0%; Score 610.8; DB 12; Length 953;
Best Local Similarity 92.1%; Pred. No. 1.7e-112;
Matches 722; Conservative 0; Mismatches 52; Indels 10; Gaps 7;
Qy 494 AAGAAGAAAAGCAATGTAATCACTATGATTTGTCACAAAACACAGAAATCTCCAAC 553
Db 1 AAGAAGAAAAGCAATGTAATCACTATGATTTGTCACAAAACACAGAAATCTCCAAC 59
Qy 554 AATTTTAAGTAATCTGATTCCTCTCTGGAGAAAATGTTACCTAATAGTTTTCCTTAT 613
Db 60 AATTTTAAGTAATCTGATTCCTCTCTGGAGAAAATGTTACCTAATAGTTTTCCTTAT 118
Qy 614 GAATGTTATTACTACTCGTATAAATCAAAATTCCTATAAATTCCTACTTAAGTCTTAAGA 673
Db 119 GAATGTTATTACTACTCGTATAAATCAAAATTCCTATAAATTCCTACTTAAGTCTTAAGA 178
Qy 674 ACTGGGTTCTCTTTCATGTTATTCATGTTTCAGAAAGGAAAACACACTTACTCTTTTA 733
Db 179 ACTGGGTTCTCTTTCATGTTATTCATGTTTCAGAAAGGAAAACACACTTACTCTTTTA 238
Qy 734 GGCAATTCCTAGTAATCTAGTAGTATCAGGATATATTTGCTTTAAATATATTTTGG 793
Db 239 GGCAATTCCTAGTAATCTAGTAGTATCAGGATATATTTGCTTTAAATATATTTTGG 295
Qy 794 TTATTTTGAATACAGACATTTGGCTCCAAATTTTCACTTTTGCACAAATAGTATGACTTTTC 853
Db 296 GTTATTTGAATACAGACATTTGGCTCCAAATTTTCACTTTTGCACAAATAGTATGACTTTTC 355
Qy 854 ACTAGAACTTCTCAACATTTGGGAACCTTCGAATATGAGCATCATATGTTAAGGCTG 913
Db 356 ACTAGAACTTCTCAACATTTGGGAACCTTCGAATATGAGCATCATATGTTAAGGCTG 415

Qy 914 TATCAATTAATGCTATGAGATACATTTGTTTCTCCCTATGCCAAAACAGGTGAACAAACGT 973
Db 416 TATCAATTAATGCTATGAGATACATTTGTTTCTCCCTATGCCAAAACAGGTGAACAAACGT 475
Qy 974 AGTTGTTTTTACTGATCTAATGTTGGCTACCTGCTGATTTTATAGTATGCACATGTCA 1033
Db 476 AGTTGTTTTTACTGATCTAATGTTGGCTACCTGCTGATTTTATAGTATGCACATGTCA 535
Qy 1034 GAAAAAGGCAAGCAAAATGGCCCTTTGTACTGAATACTTTGGCAAACTTATTTGGGTCTTC 1093
Db 536 GAAAAAGGCAAGCAAAATGGCCCTTTGTACTGAATACTTTGGCAAACTTATTTGGGTCTTC 595
Qy 1094 ATTTTCTGACAGA-CAGGATTTGACTCAATATTTGTAGAGCTTGCGTGAATCGATTACA 1152
Db 596 ATTTTCTGACAGACCAGGATTTGACTCAATATTTGTAGAGCTTGCGTGAATCGATTACA 655
Qy 1153 TGGTAGTGTGCTACCTGGTAGAAAATGGTTTTTGTAGTTATTTGACTCAGAAATTCATCTCAGGAT 1212
Db 656 TGG-AATGATGCCCTGGTAGAAA--GGTCAAAAGTTATAGACTCCGAA-TCATCTCAGGAG 711
Qy 1213 GAATCTTTTATGCTCTTTTATTTGTAAGCATATCTGAAATTTACTTTTATAGATGGTTTAA 1272
Db 712 AAATCTTTATGCTCTCTATGGAAGGCCCTTCTGATTTACTAATAACGATGGTTAAAGGCTTGC 771
Qy 1273 GAAA 1276
Db 772 TAAA 775

Search completed: April 26, 2004, 01:59:42
Job time : 4260.62 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 25, 2004, 07:15:18 ; Search time 116.382 Seconds
(without alignments)
11186.557 Million cell updates/sec

Title: US-10-051-835-25
Perfect score: 2346
Sequence: 1 gccaggggggaaaaaagag.....aaacatttgatgattacat 2346

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	581.4	24.8	675	4	US-09-702-705-1269 Sequence 1269, Ap
2	581.4	24.8	675	4	US-09-736-457-1269 Sequence 1269, Ap
3	581.4	24.8	675	4	US-09-671-325-1269 Sequence 1269, Ap
4	551.4	23.5	665	4	US-09-614-1248-1269 Sequence 1269, Ap
C 5	54.6	2.3	640681	4	US-09-790-988-1 Sequence 1, Appl
C 6	49.8	2.1	1501	4	US-09-457-037B-30 Sequence 30, Appl
C 7	49.8	2.1	1501	4	US-09-733-151-30 Sequence 30, Appl
C 8	49	2.1	694	4	US-09-457-037B-39 Sequence 39, Appl
C 9	49	2.1	694	4	US-09-733-151-39 Sequence 39, Appl
C 10	49	2.1	1279	4	US-09-457-037B-40 Sequence 40, Appl
C 11	49	2.1	1279	4	US-09-733-151-40 Sequence 40, Appl
C 12	49	2.1	6317	4	US-10-204-708-11 Sequence 11, Appl
13	48.4	2.1	112132	4	US-09-741-150-3 Sequence 3, Appl
14	48.4	2.1	112132	4	US-10-160-187-3 Sequence 2, Appl
15	47.8	2.0	10467	4	US-10-204-708-2 Patent No. 5231168-1
C 16	46	2.0	3095	6	5231168-1
C 17	45.6	1.9	580073	4	US-08-545-528D-1 Sequence 1, Appl
18	45	1.9	1689	4	US-07-991-867B-41 Sequence 41, Appl
19	45	1.9	1689	2	US-08-544-332-41 Sequence 41, Appl
20	45	1.9	1689	4	US-09-370-861A-41 Sequence 41, Appl
21	45	1.9	1947	4	US-09-370-861A-74 Sequence 74, Appl
22	45	1.9	8457	1	US-07-991-867B-1 Sequence 1, Appl
23	45	1.9	8457	2	US-08-544-332-1 Sequence 1, Appl
24	45	1.9	8457	1	US-09-370-861A-1 Sequence 1, Appl
25	45	1.9	640681	4	US-09-790-988-1 Sequence 1, Appl
C 26	44.8	1.9	4336	4	US-08-956-171E-546 Sequence 546, App
27	44.4	1.9	19124	2	US-08-487-826B-13 Sequence 13, Appl

C 28	44.2	1.9	19124	2	US-08-487-826B-13 Sequence 13, Appl
C 29	44	1.9	832	4	US-09-621-976-2813 Sequence 2813, Ap
C 30	44	1.9	6866	4	US-10-204-708-20 Sequence 20, Appl
C 31	44	1.9	20674	4	US-09-641-638-651 Sequence 651, Appl
C 32	43.8	1.9	1077	4	US-09-457-037B-24 Sequence 24, Appl
C 33	43.8	1.9	1077	4	US-09-733-151-24 Sequence 24, Appl
C 34	43.6	1.9	4140	3	US-08-894-731-2 Sequence 2, Appl
35	43.4	1.8	6113	4	US-10-204-708-14 Sequence 14, Appl
36	43.2	1.8	5844	4	US-10-204-708-89 Sequence 89, Appl
37	43	1.8	8801	4	US-10-204-708-62 Sequence 62, Appl
38	43	1.8	20674	4	US-09-641-638-651 Sequence 651, Appl
39	42.2	1.8	658	3	US-08-998-416-193 Sequence 193, App
40	42.2	1.8	676	3	US-08-998-416-1014 Sequence 1014, Ap
41	42.2	1.8	677	3	US-08-998-416-920 Sequence 920, App
42	42.2	1.8	832	4	US-09-621-976-2813 Sequence 2813, Ap
C 43	42.2	1.8	1274	4	US-09-598-401C-34 Sequence 34, Appl
44	42	1.8	5562	4	US-10-204-708-63 Sequence 63, Appl
45	42	1.8	5666	4	US-10-204-708-29 Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-09-702-705-1269
; Sequence 1269, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Pan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1269
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(675)
; OTHER INFORMATION: n = A,T,C or G
US-09-702-705-1269

Query Match	24.8%;	Score	581.4;	DB	4;	Length	675;
Best Local Similarity	97.3%;	Pred. No.	8.5e-136;				
Matches	62;	Conservative	0;	Mismatches	13;	Indels	4;
Gaps	3;						
Qy	1673	CTGAAAAGAGTGATCTCTCAATATCTCACTAAGTCTCTCACTCAAGCAGATTCT	1732				
Db	1	CTGAAAAGAGTGATCTCTCAATATCTCACTAAGTCTCTCACTCAAGCAGATTCT	60				
Qy	1733	TCACCTCGGACCTGTGATCATGAACCTAGTAGGGGATGTGTATTTTATACAAAT	1792				
Db	61	TCACCTCGGACCTGTGATCATGAACCTAGTAGGGGATGTGTATTTTATACAAAT	120				
Qy	1793	TTAATCAATGCTTACATTCATTAATCTTAAAGAGCAAACTGCATTATTTTCTGC	1852				
Db	121	TTAATCAATGCTTACATTCATTAATCTTAAAGAGCAAACTGCATTATTTTCTGC	180				
Qy	1853	ATCCACATTCACATCATATTAGAACTAAGATATTTATCTATGAAGATATAAATGGTGCAG	1912				
Db	181	ATCCACATTCACATCATATTAGAACTAAGATATTTATCTATGAAGATATAAATGGTGCAG	240				


```

RESULT 4
US-09-614-124B-1269
; Sequence 1269, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongrong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1269
; LENGTH: 665
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1..665)
; OTHER INFORMATION: n = A.T.C or G

```

```

RESULT 5
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, YOSHIAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790, 988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

```

Query Match 2.3%; Score 54.6; DB 4; Length 640681;
Best Local Similarity 45.7%; Pred. No. 0.004;
Matches 239; Conservative 0; Mismatches 274; Indels 10; Gaps 1;
QY 475 TAACTTTCTTTCTGCTGAAAGAGAAAGCAATCAATCACTATGATTAATGACAA 534
Db 574858 TATATATATATTTACTATATAGAGAAATAGTAAAAATAACCTTTAATATCTTAATATAA 574799
QY 535 ACACAGAGATTTCCACAAATTTTAACTGATCTCTCTCTGAGAGAAATTTGTTA 594
Db 574798 AAGAGAAATCTAGCTCAAGTTTCTGATTTTGTATATATCTTTTATTTGAAGATATCT 574739
QY 595 CCTAATAGTTTTCCTTATGATGTTTATTAATCACTGATTAATCAATCAATTTCTATAAAT 654
Db 574738 AATATTTTCTTATAGTAAAGTAAAGAAATTTGTTTATTTAAATATAAATAAGAAAAAT 574679
QY 655 TCCTACTTAACTTCTAAGAACTGGTT-----CTTCTCTGATGTTTATCATGTT 704
Db 574678 TATAGTAACTGATAGATATATTTTAAATAATCAATTTTATCTTTTATGTTA 574619
QY 705 CAGAAAGGAAACACACTTTTACTCTTTTAGGACAAATCCCTAGAAATCTATAGTAGTATCAG 764
Db 574618 CTAATAAGAGGATTAATTTTAAATTAATGAAATAATAACAATTTTAAATAATATACAT 574559
QY 765 GATATATTTTCTTAAATATATTTTGGTATTTTGAATACAGACATTTGGCTCCAAAT 824
Db 574558 CACTCGTTAAATTTAAATAATTTTATTAATGTTTAAATATAGAAATTTACATAATAT 574499
QY 825 TTCTCTTTGACAAATAGTATGACTTTTCTCACTAGAACTTCTCAACATTTTGGGAACTTTGC 884
Db 574498 TTTACATTTTAAATGATTTTAAATTTTAAATAATTAATTAACTTTGTCTATTTTCA 574439
QY 885 AATATGAGCATCATGCTGTTAGGCTGTATCATTTTAAAGCTATGATAGTATGATTTT 944
Db 574438 ACGTGTCTTATGCTCCGACCAAGATGCTGAGCAATGCTGATTTTAAAGCAATATC 574379
QY 945 CTCCTATGCAACAGCTGAAACAAAGCTAGTTGTTTCTTACT 987
Db 574378 TTAGTATATAAATATGCGCAGAAACGAATATTTTCTAGT 574336
RESULT 6
US-09-457-037B-30/c
; Sequence 30, Application US/09457037B
; Patent No. 6506963
; GENERAL INFORMATION:
; APPLICANT: De Both, et al.
; TITLE OF INVENTION: HYBRID WINTER OILSEED RAPE AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: 514412-2019
; CURRENT APPLICATION NUMBER: US/09/457,037B
; CURRENT FILING DATE: 2000-12-08
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 30
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1501)
; OTHER INFORMATION: "n" stands for any nucleotide a,c,g or t
; NAME/KEY: misc_feature
; LOCATION: (1)..(16)
; OTHER INFORMATION: p-GEM-T vector
; NAME/KEY: misc_feature
; LOCATION: (1458)..(1501)
; OTHER INFORMATION: p-GEM-T vector
; NAME/KEY: misc_feature
; LOCATION: (1)..(1501)
; OTHER INFORMATION: sequence comprising the 3' flanking region of RF-BN1

US-09-457-037B-30

Query Match 2.1%; Score 49.8; DB 4; Length 1501;
Best Local Similarity 46.3%; Pred. No. 0.0057;
Matches 156; Conservative 2; Mismatches 179; Indels 0; Gaps 0;
QY 1081 TTATGGGTCTTCTATTTCTTGACAGACAGGATTTGACTCAATATTTGTAGAGCTTGGTA 1140
Db 883 TGAATTCGACTTCTGGTTAAAGTAACCTAGTTTGTGATCATTTTGTCTATATTTAAATA 824
QY 1141 GAATGGATTACATGCTAGTATGATGCTAGTAAATGGTTTATTTAGTTTATTTGACTCAGAA 1200
Db 823 AACTATAGAGATTTCTACTTCTAAATAAAGCTGTAATATTTTGATCCACATTTAGCTTCTT 764
QY 1201 TCATCTCAGGATGAATCTTTTATGCTCTTTTATTTGTAAGCATATCTGAAATTTACTTTATA 1260
Db 763 TTTATATGCTAGTATGATTTTATTTATATATTTATGAGAATTTGATAAAAAATTTATA 704
QY 1261 AAGATGGTTTATGAAAGCTTTGCTAAAAATTTGGCCTAGGAATGGTAATTTCAATTTTCA 1320
Db 703 ATATTTTCCATATATAGTTTACTATAGTTTAAATTTCTAGATTTACTTTTGTGACA 644
QY 1321 GTTCCCAAGGGGTAGAAAAATAATATATGTTGTTGTTATGTTTATGTTTAAACATATATTAG 1380
Db 643 GCTACCAATATTTCAATCAAAAAATTTGTAATCTTGTATTTATATATACAAATTTCTACTTT 584
QY 1381 GTACTATCTATGATGATTTTAAATATTTTTCATTT 1417
Db 583 TTGTTATGATTTTAAATAATTTTATTTTATTTTATTTT 547
RESULT 7
US-09-733-151-30/c
; Sequence 30, Application US/09733151
; Patent No. 6563826
; GENERAL INFORMATION:
; APPLICANT: De Both, et al.
; TITLE OF INVENTION: HYBRID WINTER OILSEED RAPE AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: 514412-2019
; CURRENT APPLICATION NUMBER: US/09/733,151
; CURRENT FILING DATE: 2000-12-08
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 30
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1501)
; OTHER INFORMATION: "n" stands for any nucleotide
; OTHER INFORMATION: "b" stands for either nucleotide g or c
; OTHER INFORMATION: "w" stands for nucleotide a or t/u
; OTHER INFORMATION: sequence comprising the 3' flanking region of RF-BN1.

US-09-733-151-30
Query Match 2.1%; Score 49.8; DB 4; Length 1501;
Best Local Similarity 46.3%; Pred. No. 0.0057;
Matches 156; Conservative 2; Mismatches 179; Indels 0; Gaps 0;
QY 1081 TTATGGGTCTTCTATTTCTGACAGACAGGATTTGACTCAATATTTGTAGAGCTTGGTA 1140
Db 883 TGAATTCGACTTCTGGTTAAAGTAACCTAGTTTGTGATCAATTTTGTCTATATTTAAATA 824
QY 1141 GAATGGATTACATGCTAGTATGATGCTAGTAAATGGTTTATTTAGTTTATTTGACTCAGAA 1200
Db 823 AACTATAGAGATTTCTACTTCTAAATAAAGCTGTAATATTTTGATCCACATTTAGCTTCTT 764
QY 1201 TCATCTCAGGATGAATCTTTTATGCTCTTTTATTTGTAAGCATATCTGAAATTTACTTTATA 1260
Db 763 TTTATATGCTATGATGATTTTATTTATATATTTATGAGAATTTGATAAAAAATTTATA 704

QY 1261 AAGATGGTTTGAAGAAGCTTTGCTTAAAAATTTGGCCCTAGGAATGGTAACTTTCATTTTCA 1320
DB 703 ATATTTTCCATATATAGTTTACTATAGTTAAATTTAAATTTCTAAGATTTACTTTTGGACA 644
QY 1321 GTTGGCAAGGGGTAGAAAAAATAATATGTGTGTATATGTTTATGTTTAAACATATATTAG 1380
DB 643 GCTACCAATATTTCAATCAAAAATTTGTATCTTGTATTTATATATACAAATTCATCTTT 584
QY 1381 GTACTATCTATGAATGATTTTAAATATTTTTCATATT 1417
DB 583 TTGTTTCATGATTTKAAAAATTTTAGTTTATAAAATTTT 547

RESULT 8

US-09-457-037B-39/c
; Sequence 39, Application US/09457037B
; Patent No. 6506963
; GENERAL INFORMATION:
; APPLICANT: De Both, et al.
; TITLE OF INVENTION: HYBRID WINTER OILSEED RAPE AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: 514412-2019
; CURRENT APPLICATION NUMBER: US/09/457,037B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/457,037
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 694
; TYPE: DNA
; ORGANISM: Brassica napus
; NAME/KEY: misc_feature
; LOCATION: (1)..(694)
; OTHER INFORMATION: sequence comprising the 5' flanking region of RF-BN1 in WOSR
US-09-457-037B-39

Query Match 2.1%; Score 49; DB 4; Length 694;
Best Local Similarity 46.6%; Pred. No. 0.0067;
Matches 157; Conservative 0; Mismatches 180; Indels 0; Gaps 0;
QY 1081 TTATTGGGTCTTCATTTCTGACAGACAGAGATTGACTCAATATTTGTAGAGCTTGGCTA 1140
DB 462 TGAATGGACTTCTGGTTAAAGTAACTAGTTTGTGATCAATTTTGTCTATATTTAAATA 403
QY 1141 GAATGATTACATGGTAGTGATGCACTGGTAGAAATGGTTTTAGTTATTCGACTCAGAAAT 1200
DB 402 AAACATATAAGATTTCTACTTCTAAAAAACTGTAATAATTTTGTATCCATTTCTTCTT 343
QY 1201 TCATCTCAGGATGAATCTTTTATGCTTTTATTTTATTTGTAAGCATATCTGAAATTTTATA 1260
DB 342 TTTATATGCTATGAGTTAGATTTATTTTATATATTTTATGAGNAATTTGATAAAATTTATA 283
QY 1261 AAGATGGTTTGAAGAAGCTTTCTTAAAAATTTGGCCCTAGGAATGGTAACTTTCATTTTCA 1320
DB 282 ATATTTTCAATATATAGTTTACTATAGTTAAATTTAAATTTCTAAGATTTACTTTTGGACA 223
QY 1321 GTTGGCAAGGGGTAGAAAAAATAATATGTGTGTATGTTTATGTTTAAACATATATTAG 1380
DB 222 GCTACCAATATTTCAATCAAAAATTTGTATCTTGTATTTATATATACAAATTCATCTTT 163
QY 1381 GTACTATCTATGAATGATTTTAAATATTTTTCATATT 1417
DB 162 TTGTTTCATGATTTAAAAATTTTAGTTTATAAAATTTT 126

RESULT 9

US-09-457-037B-40
; Sequence 40, Application US/09457037B
; Patent No. 6506963
; GENERAL INFORMATION:
; APPLICANT: De Both, et al.
; TITLE OF INVENTION: HYBRID WINTER OILSEED RAPE AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: 514412-2019
; CURRENT APPLICATION NUMBER: US/09/457,037B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/457,037
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 40
; LENGTH: 1279
; TYPE: DNA
; ORGANISM: Brassica napus
; NAME/KEY: misc_feature
; LOCATION: (1)..(1279)
; OTHER INFORMATION: sequence comprising the 3' flanking region of RF-BN1 in WOSR
US-09-457-037B-40

Query Match 2.1%; Score 49; DB 4; Length 1279;
Best Local Similarity 46.6%; Pred. No. 0.0085;

; TITLE OF INVENTION: HYBRID WINTER OILSEED RAPE AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: 514412-2019
; CURRENT APPLICATION NUMBER: US/09/733,151
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/457,037
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 694
; TYPE: DNA
; ORGANISM: Brassica napus
; NAME/KEY: misc_feature
; LOCATION: (1)..(694)
; OTHER INFORMATION: sequence comprising the 5' flanking region of RF-BN1 in WOSR
US-09-733-151-39

Query Match 2.1%; Score 49; DB 4; Length 694;
Best Local Similarity 46.6%; Pred. No. 0.0067;
Matches 157; Conservative 0; Mismatches 180; Indels 0; Gaps 0;
QY 1081 TTATTGGGTCTTCATTTCTGACAGACAGAGATTGACTCAATATTTGTAGAGCTTGGCTA 1140
DB 462 TGAATGGACTTCTGGTTAAAGTAACTAGTTTGTGATCAATTTTGTCTATATTTAAATA 403
QY 1141 GAATGATTACATGGTAGTGATGCACTGGTAGAAATGGTTTTAGTTATTCGACTCAGAAAT 1200
DB 402 AAACATATAAGATTTCTACTTCTAAAAAACTGTAATAATTTTGTATCCATTTCTTCTT 343
QY 1201 TCATCTCAGGATGAATCTTTTATGCTTTTATTTTATTTGTAAGCATATCTGAAATTTTATA 1260
DB 342 TTTATATGCTATGAGTTAGATTTATTTTATATATTTTATGAGNAATTTGATAAAATTTATA 283
QY 1261 AAGATGGTTTGAAGAAGCTTTCTTAAAAATTTGGCCCTAGGAATGGTAACTTTCATTTTCA 1320
DB 282 ATATTTTCAATATATAGTTTACTATAGTTAAATTTAAATTTCTAAGATTTACTTTTGGACA 223
QY 1321 GTTGGCAAGGGGTAGAAAAAATAATATGTGTGTATGTTTATGTTTAAACATATATTAG 1380
DB 222 GCTACCAATATTTCAATCAAAAATTTGTATCTTGTATTTATATATACAAATTCATCTTT 163
QY 1381 GTACTATCTATGAATGATTTTAAATATTTTTCATATT 1417
DB 162 TTGTTTCATGATTTAAAAATTTTAGTTTATAAAATTTT 126

RESULT 10

US-09-457-037B-40/c
; Sequence 40, Application US/09457037B
; Patent No. 6506963
; GENERAL INFORMATION:
; APPLICANT: De Both, et al.
; TITLE OF INVENTION: HYBRID WINTER OILSEED RAPE AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: 514412-2019
; CURRENT APPLICATION NUMBER: US/09/457,037B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/457,037
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 40
; LENGTH: 1279
; TYPE: DNA
; ORGANISM: Brassica napus
; NAME/KEY: misc_feature
; LOCATION: (1)..(1279)
; OTHER INFORMATION: sequence comprising the 3' flanking region of RF-BN1 in WOSR
US-09-457-037B-40

Query Match 2.1%; Score 49; DB 4; Length 1279;
Best Local Similarity 46.6%; Pred. No. 0.0085;

Matches 157; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

QY 1081 TTATGGGCTTCATTTTCGACAGAGATTGACTCAATATTGTAGAGCTTGCCTA 1140
DB 740 TGAATTGGACCTTCGTTAAAGTAAGTACTAGTTTGTGATCATTTTGGCTATATTTAAATA 681

QY 1141 GAATGATTACATGCTAGTATGATGCTAGTAAATGGTTTTAGTTATGACTCAGAA 1200
DB 680 AACTATAAGATCTTACTTCTTAAATACTGTAATATTTTGATCCAGATTAGTCTTCT 621

QY 1201 TCATCTCAGATGAATCTTTTANGCTTTTATTTGTAAGCATATCTGAAATTTACTTTATA 1260
DB 620 TTATATGCTATGATGATTAGTTATTTTATATTTATGAGAAATTTGATAAAAAATTTATA 561

QY 1261 AAGATGTTTTAGAAAGCTTTGCTAAATAATTTGGCTAGGAATGTAATCTTCAATTTCA 1320
DB 560 ATATTTCAATATATAGTTTACTATAGTTAAATTAATTTCTAGATTTACTTTTTTGACA 501

QY 1321 GTTGCCAGGGGAGAAAAATATATATGTTGTTGTTATGTTTATGTTAATATTTATTAG 1380
DB 500 GCTACCAATATTTCAATCAAAATTTGTATCTTTGTATTTATATTAACAAATTCATCTT 441

QY 1381 GTACTATCTATGATGATTTAATATATTTTTCATATT 1417
DB 440 TTGTCATGATTTAAAAATTTTGTATTTAATAAATTTT 404

RESULT 11
US-09-733-151-40/c
; Sequence 40, Application US/09733151
; Patent No. 6563026
; GENERAL INFORMATION:
; APPLICANT: De Both, et al.
; TITLE OF INVENTION: HYBRID WINTER OILSEED RAPE AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: 514412-2019
; CURRENT APPLICATION NUMBER: US/09/733,151
; CURRENT FILING DATE: 2000-12-08
; PRIOR FILING DATE: 09/457,037
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 40
; LENGTH: 1279
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1279)
; OTHER INFORMATION: sequence comprising the 3' flanking region of RP-BN1 in WOSR
US-09-733-151-40

Query Match 2.1%; Score 49; DB 4; Length 1279;
Best Local Similarity 46.6%; Pred. No. 0.0085;
Matches 157; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

QY 1081 TTATGGGCTTCATTTTCGACAGAGATTGACTCAATATTGTAGAGCTTGCCTA 1140
DB 740 TGAATTGGACCTTCGTTAAAGTAAGTACTAGTTTGTGATCATTTTGGCTATATTTAAATA 681

QY 1141 GAATGATTACATGCTAGTATGATGCTAGTAAATGGTTTTAGTTATGACTCAGAA 1200
DB 680 AACTATAAGATCTTACTTCTTAAATACTGTAATATTTTGATCCAGATTAGTCTTCT 621

QY 1201 TCATCTCAGGATGAATCTTTTATGCTTTTATTTGTAAGCATATCTGAATTTACTTTATA 1260
DB 620 TTATATGCTATGATGATTAGTTATTTTATATATTTATGAGAAATTTGATAAAAAATTTATA 561

QY 1261 AAGATGTTTTAGAAAGCTTTGCTAAATAATTTGGCTAGGAATGTAATCTTCAATTTCA 1320
DB 560 ATATTTCAATATATAGTTTACTATAGTTAAATTAATTTCTAGATTTACTTTTTTGACA 501

QY 1321 GTTGCCAGGGGAGAAAAATATATGTTGTTGTTATGTTTATGTTAATATTTATTAG 1380

DB 500 GCTACCAATATTTCAATCAAAATTTTGTATCTTTGTATTTATATTAACAAATTCATCTT 441

QY 1381 GTACTATCTATGATGATTTAATATATTTTTCATATT 1417

DB 440 TTGTCATGATTTAAAAATTTTGTATTTAATAAATTTT 404

RESULT 12
US-10-204-708-11
; Sequence 11, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 11
; LENGTH: 6317
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-11

Query Match 2.1%; Score 49; DB 4; Length 6317;
Best Local Similarity 45.4%; Pred. No. 0.016;
Matches 213; Conservative 0; Mismatches 255; Indels 1; Gaps 1;

QY 551 AACCAATTTTAAAGTAATCTGATCTCTCTTTGGAGAAAATTTGTTACCTAAATAGTTTTTCCT 610
DB 3790 AAATATAGTAAAAATTTTATATGATTTTAAATAGTTTATTTATTTATTTATTTAT 3849

QY 611 TATGAATGTTTATTTACTACTGTTATTAATCAATCTTATATAATTTCTTACTTAAGCTTA 670
DB 3850 GTTTTAAATATTTGTTTGAATTTTAAATAGGAATTTTAAATATAATATTTTAAATATT 3909

QY 671 AGAACTGGGTTCTTCTTCTGATTTTATTCATGTTCA-GAAAGGAAAACACACTTTTACTCT 729
DB 3910 TTATAGTATTTTATTTTGTAGATAAGAGATTAAGGATAGATAATTTAGTTGTTAT 3969

QY 730 TTTAGGCAATCTCTAGAACTATAGTAGTATACGAGTATATTTTGTCTTAAATATATT 789
DB 3970 AATGCTAGAAATTAATGGAATGGAGTGAATTTAGGTTTAAATTTAGTAGAGATTATAGT 4029

QY 790 TTGGTTATTTTGAATACAGACATTTGGCTCCAAATTTTTCATCTTTGCCAATAGTATGACT 849
DB 4030 TTTTATTTTATTTTAAATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4089

QY 850 TTTCACTAGAACTTCTCAACATTTGGGAATTTTGGCAATATGAGCATCATATGTTGTAAG 909
DB 4090 ATTTGTTTAAAGTATTTATATGTTTAAAAATTTTATTTAGTTTAAATTTTGAATGATTTT 4149

QY 910 GCTGTATCATTTAATGCTATGAGATACATTTGTTTCTCCCTATGCCAAACAGGTGAACAA 969
DB 4150 ATTTAATTTAGAAATTAATTTTGTGTTAGTTTAAATTTTAAATTAATAATAATAA 4209

QY 970 ACGTAGTTGTTTTTACTGATCTATAATGTTGGCTACCTGTGATTTTAT 1018

Db 4210 AATAAGATATAGTTTATAGATTATATATGTTAGTGGAGAGTTT 4258

RESULT 13
US-09-741-150-3
; Sequence 3, Application US/09741150
; Patent No. 6436689
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: US\$ THEREOF
; FILE REFERENCE: CL000968
; CURRENT APPLICATION NUMBER: US/09/741,150
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 112132
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(112132)
; OTHER INFORMATION: n = A,T,C or G
US-09-741-150-3

Query Match 2.1%; Score 48.4; DB 4; Length 112132;
Best Local Similarity 55.3%; Pred. No. 0.07;
Matches 94; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 720 ACTTACTCTTTAGGACAATTCCTAGATCTATAGTATCAGGATATATTTGCTTT 779
DB 26831 ACTTCTACCTAGAGCTCTGACTAAATCAATAGTACTATTAATTAATATAGTTT 26890

QY 780 AAAATATATTTGGTATTTTGAATCAGACATTTGGCTCCAAATTTTCATCTTTGCACAA 839
DB 26891 ACTGCTTAGCTAGGTTTGGGGGACTAGCTTGGGAACCAATACCATCTCAGGCCAT 26950

QY 840 TAGTATGACTTTTCACTAGAACCTTCTCAACATTTGGGAACCTTTGCCAATA 889
DB 26951 TTTTTCCTTTATGAATATATCCTTAGCAAAATTTCTAAATTAATTAATAA 27000

LENGTH: 10467
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-2

Query Match 2.0%; Score 47.8; DB 4; Length 10467;
Best Local Similarity 45.6%; Pred. No. 0.039;
Matches 169; Conservative 0; Mismatches 202; Indels 0; Gaps 0;

QY 289 TTTCGACTGTCTTATTTTCAAGATTTCTCAATTAATTTTCCCTAGAAATTTCTCCTTCA 348
DB 7895 TTTAGTGTTTTATATTAAGTTTATTTTTCGTTAGTATTTATAGTTATATTTAT 7954

QY 349 TTCAAAGTACAAATCTTTGAAGAAATGAACAGATTTCCCATGAAATGATGTCTCA 408
DB 7955 ATTTTGTGATTAATAATATTTGAATTTTAAAAATTTATATTTTAAATTTTAAAA 8014

QY 409 TACTCGACTAGAACGATCTATGTTAATGACGTGTATATGAATTTTCAAGTACTAC 468
DB 8015 TGTTATGTTTAAATGTTTGTGTTTGTAAATAGTATTTTGAATTAATTAATTTTG 8074

QY 469 CCCAATTAACCTTTCTTATTCCTCTGAAAGAAAGCAATGTAATCACTATGATTT 528
DB 8075 GTTTGATAATTTTGTGTTGAGTTTATAATTTTATTTGCGAATATCGAAATAT 8134

QY 529 GCACAAACACCAAGATTTCTCAACAAATTTTAAATGATCTCTCTTCTTGGAGAAA 588
DB 8135 TTGGATATATTTATATTTTCTGTTTATTTTATTTTATTTTGTGTTTCTTACGTTGTTT 8194

QY 589 TTGTACCTAATAGTTTTTCCCTTATGAATGTTTATTACTACTGTTAATAATCAAAATTTCTA 648

Best Local Similarity 55.3%; Pred. No. 0.07;
Matches 94; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 720 ACTTACTCTTTAGGACAATTCCTAGAAATCTATAGTATCAGGATATATTTGCTTT 779
DB 26831 ACTTCTACCTAGGACTCTGACTAAATCAATAGTACTATTAATTAATATAGTTT 26890

QY 780 AAAATATATTTGGTATTTTGAATCAGACATTTGGCTCCAAATTTTCATCTTTGCACAA 839
DB 26891 ACTGCTTAGCTAGGTTTGGGGGACTAGCTTGGGAACCAATACCATCTCAGGCCAT 26950

QY 840 TAGTATGACTTTTCACTAGAACCTTCTCAACATTTGGGAACCTTTGCCAATA 889
DB 26951 TTTTTCCTTTATGAATATATCCTTAGCAAAATTTCTAAATTAATTAATAA 27000

RESULT 15
US-10-204-708-2
; Sequence 2, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 2
; LENGTH: 10467
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-2

Query Match 2.0%; Score 47.8; DB 4; Length 10467;
Best Local Similarity 45.6%; Pred. No. 0.039;
Matches 169; Conservative 0; Mismatches 202; Indels 0; Gaps 0;

QY 289 TTTCGACTGTCTTATTTTCAAGATTTCTCAATTAATTTTCCCTAGAAATTTCTCCTTCA 348
DB 7895 TTTAGTGTTTTATATTAAGTTTATTTTTCGTTAGTATTTATAGTTATATTTAT 7954

QY 349 TTCAAAGTACAAATCTTTGAAGAAATGAACAGATTTCCCATGAAATGATGTCTCA 408
DB 7955 ATTTTGTGATTAATAATATTTGAATTTTAAAAATTTATATTTTAAATTTTAAAA 8014

QY 409 TACTCGACTAGAACGATCTATGTTAATGACGTGTATATGAATTTTCAAGTACTAC 468
DB 8015 TGTTATGTTTAAATGTTTGTGTTTGTAAATAGTATTTTGAATTAATTAATTTTG 8074

QY 469 CCCAATTAACCTTTCTTATTCCTCTGAAAGAAAGCAATGTAATCACTATGATTT 528
DB 8075 GTTTGATAATTTTGTGTTGAGTTTATAATTTTATTTTATTTGCGAATATCGAAATAT 8134

QY 529 GCACAAACACCAAGATTTCTCAACAAATTTTAAATGATCTCTCTTCTTGGAGAAA 588
DB 8135 TTGGATATATTTATATTTTCTGTTTATTTTATTTTATTTTGTGTTTCTTACGTTGTTT 8194

QY 589 TTGTACCTAATAGTTTTTCCCTTATGAATGTTTATTACTACTGTTAATAATCAAAATTTCTA 648

Db 8195 ATTTTAAATTTTTCGANTGATTTTATTTTATTTTAAATTTTATAGTT 8254
Qy 649 TAAATTCCTA 659
Db 8255 TCAAAAGTTGTA 8265

Search completed: April 26, 2004, 02:11:20
Job time : 120.382 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2004, 23:47:42 ; Search time 623.343 Seconds

(without alignments)
15988.422 Million cell updates/sec

Title: US-10-051-835-25

Perfect score: 2346

Sequence: 1 gccagaggggaaaaaagag.....aaacatttgatgttaccat 2346

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 337863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04:*

- 1: Geneseqm1980as:*
- 2: Geneseqm1990as:*
- 3: Geneseqm2000as:*
- 4: Geneseqm2000as:*
- 5: Geneseqm2000as:*
- 6: Geneseqm2002as:*
- 7: Geneseqm2003as:*
- 8: Geneseqm2003bs:*
- 9: Geneseqm2003cs:*
- 10: Geneseqm2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2346	100.0	2346	8 ACH04269	ACH04269 Human CDN
2	2334	99.5	2350	9 ADE25669	ADE25669 Human CDN
3	2304.4	98.2	2321	6 AAS94792	AAS94792 Human DNA
4	1599.4	68.2	1626	3 AAS98844	AAS98844 Human dif
5	1599.4	68.2	1626	3 AAS98125	AAS98125 Human dif
6	1599.4	68.2	1626	6 AAD29955	AA29955 Human pha
7	603.2	25.3	665	4 AAB34083	AAB34083 Human col
8	593	25.3	603	6 ABZ08185	ABZ08185 Human leu
9	584.8	24.9	608	4 AAL22443	AAL22443 Human bre
10	582.8	24.8	586	6 ABK27756	ABK27756 Human col
11	581.4	24.8	675	6 ABK39231	ABK39231 DNA encod
12	581.4	24.8	675	7 ACA11560	ACA11560 Human lun
13	581.4	24.8	675	7 ACA02746	ACA02746 Lung can
14	568	24.2	586	6 ABU37834	ABU37834 Human col
15	546.2	23.3	783	4 AAL13574	AAL13574 Human bre
16	497.4	21.2	499	6 ABV87830	ABV87830 Human col
17	483	20.6	483	6 ABV87118	ABV87118 Human col
18	468	19.9	482	4 AAB36474	AAB36474 Human col
19	435	18.5	435	6 ABV87338	ABV87338 Human col
20	401.8	17.1	2000	7 ABF17369	ABF17369 Human SLC
21	323	13.8	323	6 ABU65495	ABU65495 Lung can
22	323	13.8	323	6 ABK64193	ABK64193 Human ben
23	323	13.8	323	6 ABN93962	ABN93962 Gene #460

C	24	322.2	13.7	324	6	ABL38369	ABL38369 Human col
C	25	321.2	13.7	324	6	ABL37940	ABL37940 Human col
C	26	321	13.7	321	4	AAD23574	AAD23574 Human lun
C	27	321	13.7	321	9	ADD66862	ADD66862 Human lun
C	28	321	13.7	321	9	ADE88116	ADE88116 Human lun
C	29	319.6	13.6	324	6	ABL38210	ABL38210 Human col
C	30	273.2	11.6	277	6	ABL38570	ABL38570 Human col
C	31	258.4	11.0	284	2	AAT25840	AAT25840 Human gen
C	32	243.6	10.4	247	6	ABL37563	ABL37563 Human col
C	33	242.6	10.3	247	6	ABL38537	ABL38537 Human col
C	34	242.6	10.3	247	6	ABL38589	ABL38589 Human col
C	35	242.4	10.3	244	6	ABK27670	ABK27670 Human col
C	36	229.6	9.8	248	6	ABL37833	ABL37833 Human col
C	37	64	2.7	309	6	ABL38438	ABL38438 Human col
C	38	61	2.6	8056	7	ABZ10246	ABZ10246 Haematopo
C	39	59.2	2.5	700	4	AAB93026	AAB93026 Human inf
C	40	58.8	2.5	5982	6	ABL34049	ABL34049 Human imm
C	41	57.4	2.4	8056	7	ABZ10246	ABZ10246 Haematopo
C	42	57.4	2.4	8056	7	ABZ10100	ABZ10100 Haematopo
C	43	57.4	2.4	18218	6	ABL33949	ABL33949 Human imm
C	44	57.2	2.4	13584	6	ABL32615	ABL32615 Human imm
C	45	57.2	2.4	17538	6	ABL33157	ABL33157 Human imm

ALIGNMENTS

RESULT 1

ACH04269

ID ACH04269 standard; cDNA; 2346 BP.

XX ACH04269;

DT 26-SEP-2003 (first entry)

XX Human cDNA differentially expressed in lung cancer #474.

XX Gene therapy; emphysema; ss; gene; chronic obstructive pulmonary disease;
XX respiratory disorder; lung cancer; asthma; human.

OS Homo sapiens.

XX US2003065157-A1.

PD 03-APR-2003.

PF 04-APR-2002; 2002US-00116802.

PR 04-APR-2001; 2001US-0281593P.

XX (LASE/) LASEK A W.

XX Lasek AW;

XX WPI; 2003-540803/51.

XX New combination comprising cDNAs that are differentially expressed in
PT respiratory disorders, useful for diagnosing or treating respiratory
PT disorders e.g., lung cancer, chronic obstructive pulmonary disease,
PT emphysema or asthma.

PS Claim 1; Page; 39pp; English.

XX The invention relates to a combination comprising cDNAs or their
CC complements that are differentially expressed in respiratory disorder.
CC The combination is useful for preparing a composition for diagnosing or
CC treating respiratory disorders e.g., lung cancer, chronic obstructive
CC pulmonary disease, emphysema or asthma. The present sequence represents
CC human cDNA differentially expressed during lung cancer

SQ Sequence 2346 BP; 712 A; 361 C; 428 G; 845 T; 0 U; 0 Other;

Query Match 100.0%; Score 2346; DB 8; Length 2346;

Best Local Similarity 100.0%; Pred. No. 0;				Matches 2346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	GCAGAGGGGAAAAAGAGTAATGCAAGGATCTCTTTTGCAGTGGTGAATTTT	60				
Dd	1	GCAGAGGGGAAAAAGAGTAATGCAAGGATCTCTTTTGCAGTGGTGAATTTT	60				
Qy	61	GAGTACCTTTGTGTGACAGGGTATTATTACAGCATCTGTGGGAAAAACCTATTAGGCCCTT	120				
Dd	61	GAGTACCTTTGTGTGACAGGGTATTATTACAGCATCTGTGGGAAAAACCTATTAGGCCCTT	120				
Qy	121	TGCATGTTTAAGCTGTATTAATTTGGTGTGAGTGGTCTGACTTAATGTGTATTAT	180				
Dd	121	TGCATGTTTAAGCTGTATTAATTTGGTGTGAGTGGTCTGACTTAATGTGTATTAT	180				
Qy	181	AAAAATTAGACATCAAAATTTTCTACTAACTAATCTTTATAGATGCATATCTTGGAGCAC	240				
Dd	181	AAAAATTAGACATCAAAATTTTCTACTAACTAATCTTTATAGATGCATATCTTGGAGCAC	240				
Qy	241	AGTCATATCACACTGGGAGGCAATGCAATGTGGTTACCTGGTCCTAGGTTGAACTGTCT	300				
Dd	241	AGTCATATCACACTGGGAGGCAATGCAATGTGGTTACCTGGTCCTAGGTTGAACTGTCT	300				
Qy	301	TATTTCAAAGATTTCTGAAATTAATTTTCCCTAGAAATTTCTCTTCAATTCCAAAGTACA	360				
Dd	301	TATTTCAAAGATTTCTGAAATTAATTTTCCCTAGAAATTTCTCTTCAATTCCAAAGTACA	360				
Qy	361	AACATATCTTTGAAGATGAAACAGATGTTCCCATGAAATGTATGCTCATCTCGACTAGA	420				
Dd	361	AACATATCTTTGAAGATGAAACAGATGTTCCCATGAAATGTATGCTCATCTCGACTAGA	420				
Qy	421	AACGATCTATGTTAAATGATGCTGTATATGAATTTTCAAGTACTTACCCCAAAATACTT	480				
Dd	421	AACGATCTATGTTAAATGATGCTGTATATGAATTTTCAAGTACTTACCCCAAAATACTT	480				
Qy	481	TCATTATGCTCTGAAGAAGAAAGCAATGTAAATCATATGATATTGTCACAAACCAACC	540				
Dd	481	TCATTATGCTCTGAAGAAGAAAGCAATGTAAATCATATGATATTGTCACAAACCAACC	540				
Qy	541	AGAACTCTCAACAAATTTAAGTAATCTGATCTCTCTTTGGAGAAAATTTGTACTTAAT	600				
Dd	541	AGAACTCTCAACAAATTTAAGTAATCTGATCTCTCTTTGGAGAAAATTTGTACTTAAT	600				
Qy	601	AGTTTTTCCCTATGAATGTTATTACTAGTGGTATAAATCAAAATTTCTATAAATTTCTTAC	660				
Dd	601	AGTTTTTCCCTATGAATGTTATTACTAGTGGTATAAATCAAAATTTCTATAAATTTCTTAC	660				
Qy	661	TTAAGTCTTAAGAACTGGGTCTCTCTTGAATGCTTATTCATGTTCAAGAAAGGAAACAACA	720				
Dd	661	TTAAGTCTTAAGAACTGGGTCTCTCTTGAATGCTTATTCATGTTCAAGAAAGGAAACAACA	720				
Qy	721	CTTTACTCTTTTAGGACAAATTCCTAGAACTATAGTAGTATCAGGATATATTGCTTTA	780				
Dd	721	CTTTACTCTTTTAGGACAAATTCCTAGAACTATAGTAGTATCAGGATATATTGCTTTA	780				
Qy	781	AAATATATTTTGGTATTTTGAATACAGACATTTGGCTCCAAATTTTCACTTTTGCACAAT	840				
Dd	781	AAATATATTTTGGTATTTTGAATACAGACATTTGGCTCCAAATTTTCACTTTTGCACAAT	840				
Qy	841	AGTATGACTTTTCACTAGAACTTTCTCAACATTTTGGGAACTTTGCAAAATATGAGCATCAT	900				
Dd	841	AGTATGACTTTTCACTAGAACTTTCTCAACATTTTGGGAACTTTGCAAAATATGAGCATCAT	900				
Qy	901	TGCTTAAAGGCTGTATCATTTAATGCTATGAGATACATTTGTTTCTCCCTATGCCAAACA	960				
Dd	901	TGCTTAAAGGCTGTATCATTTAATGCTATGAGATACATTTGTTTCTCCCTATGCCAAACA	960				
Qy	961	GGTGACAAACGTAAGTGTGTTTTTACTGATACATAATGTTGGCTACTGTGATTTTATAG	1020				
Dd	961	GGTGACAAACGTAAGTGTGTTTTTACTGATACATAATGTTGGCTACTGTGATTTTATAG	1020				
Qy	1021	TATGCATGTCAGAAAAGGCAAGACAAATGGCTCTTGTACTGAATATCTTCGGCAAC	1080				
Dd	1021	TATGCATGTCAGAAAAGGCAAGACAAATGGCTCTTGTACTGAATATCTTCGGCAAC	1080				

1021	TATGCATGTCAGAAAAGGCAAGACAAATGGCTCTTGTACTGAATATCTTCGGCAAC	1080
1081	TTATTGGGCTTTCATTTTCTGACAGACAGGATTTGATCAATATTTTGTAGAGCTTCGTA	1140
1081	TTATTGGGCTTTCATTTTCTGACAGACAGGATTTGATCAATATTTTGTAGAGCTTCGTA	1140
1141	GAATGATTCATGCTAGTGTAGTGCACCTGCTAGAAATGGTTTTTGTATTGACTCAGAAT	1200
1141	GAATGATTCATGCTAGTGTAGTGCACCTGCTAGAAATGGTTTTTGTATTGACTCAGAAT	1200
1201	TCATCTCAGGATGAATCTTTTATGCTCTTTTATTTGTAAGCATATCTGAATTTACTTTATA	1260
1201	TCATCTCAGGATGAATCTTTTATGCTCTTTTATTTGTAAGCATATCTGAATTTACTTTATA	1260
1261	AGAATGGTTTTGAAGCTTTGCTTAAATTTTGGCTAGGAATGGTAACTTCATTTTCA	1320
1261	AGAATGGTTTTGAAGCTTTGCTTAAATTTTGGCTAGGAATGGTAACTTCATTTTCA	1320
1321	GTTCGCAAGGGGTAGAAAATAATATATGTGTGTTTATGTTATGTTAAATATTTATAG	1380
1321	GTTCGCAAGGGGTAGAAAATAATATATGTGTGTTTATGTTATGTTAAATATTTATAG	1380
1381	GTACTATCTATGAATGTATTTAAATATTTTTCATATCTCTGTGACAGCATTTATAATTTG	1440
1381	GTACTATCTATGAATGTATTTAAATATTTTTCATATCTCTGTGACAGCATTTATAATTTG	1440
1441	CAACAAGTGGAGTCCATTTAGCCAGTGGGAAAGTCTTGGAACTCAGGTTTACCCTTTGAAG	1500
1441	CAACAAGTGGAGTCCATTTAGCCAGTGGGAAAGTCTTGGAACTCAGGTTTACCCTTTGAAG	1500
1501	GATATGCTGGCAGCATCTCTTTGATCTGTGCTTAACTGTAAATTTATAGCAGCTAAA	1560
1501	GATATGCTGGCAGCATCTCTTTGATCTGTGCTTAACTGTAAATTTATAGCAGCTAAA	1560
1561	TCCCTAACTTTGGATCTGGAATGCAATTTATGACCTTTGACCATTTCCCAAGATTTTCAGG	1620
1561	TCCCTAACTTTGGATCTGGAATGCAATTTATGACCTTTGACCATTTCCCAAGATTTTCAGG	1620
1621	GGCATCTGGTGGTCTAGTGTGAAACACACAGAAACAGAGAGATCCAGCTGAAAAA	1680
1621	GGCATCTGGTGGTCTAGTGTGAAACACACAGAAACAGAGAGATCCAGCTGAAAAA	1680
1681	GAGTGATCTTCAATATCTTAACCTGCTCAACTCAAGCAGAGATTTCTTCACTCTG	1740
1681	GAGTGATCTTCAATATCTTAACCTGCTCAACTCAAGCAGAGATTTCTTCACTCTG	1740
1741	GCATCTGATCATGAAACTTTAGTAGAGGGGATTTGTGTATTTTATACAAATTTAATACA	1800
1741	GCATCTGATCATGAAACTTTAGTAGAGGGGATTTGTGTATTTTATACAAATTTAATACA	1800
1801	ATGCTTTACATTTGATAAATTTCTTAAAGAGCAAACTGCAATTTTATTTCTGCAATCCACAT	1860
1801	ATGCTTTACATTTGATAAATTTCTTAAAGAGCAAACTGCAATTTTATTTCTGCAATCCACAT	1860
1861	TCCCAATCATATTTAGAACTTAAGATATTTATCTATGAGATATAAATGGTGGCAGAGACTT	1920
1861	TCCCAATCATATTTAGAACTTAAGATATTTATCTATGAGATATAAATGGTGGCAGAGACTT	1920
1921	TCATCTGTGGATTCGGTGTGTTTCTTAGGCTTCCCTAGCACTGATGCTTCAAGCATGTG	1980
1921	TCATCTGTGGATTCGGTGTGTTTCTTAGGCTTCCCTAGCACTGATGCTTCAAGCATGTG	1980
1981	ATATGTGAAATATAAATGGATTTCTTATAGCTAAATGAGTTCCCTCTGGGAGAGTTCTG	2040
1981	ATATGTGAAATATAAATGGATTTCTTATAGCTAAATGAGTTCCCTCTGGGAGAGTTCTG	2040
2041	GTACTGTGAATCACAAATGCCAGATGGTCTTTATGAGCTATTTGTCTAGTAAGTGTGAAGA	2100
2041	GTACTGTGAATCACAAATGCCAGATGGTCTTTATGAGCTATTTGTCTAGTAAGTGTGAAGA	2100
2101	TGCTATGAAGTAAGTGTGTTTTTCTTATGTTTTCATCTTATGGAACCTCTTGTGATGCTTTTG	2160
2101	TGCTATGAAGTAAGTGTGTTTTTCTTATGTTTTCATCTTATGGAACCTCTTGTGATGCTTTTG	2160

Db 840 AGTATGACTTTTCACATAGAACCTTCTCAACATTTGGGAACCTTGGCAATATGAGCATCAT 899
Qy 901 TGTGTTAAGGCTGTATCATTTAATGCTATGAGATACATTTGTTTCTCCCTATGCOAAACA 960
Db 900 TGTGTTAAGGCTGTATCATTTAATGCTATGAGATACATTTGTTTCTCCCTATGCOAAACA 959
Qy 961 GGTGACAAACCTAGTGTGTTTTTACTGATCACTAAATGTTGGCTACCTGTGATTTTATAG 1020
Db 960 GGTGACAAACCTAGTGTGTTTTTACTGATCACTAAATGTTGGCTACCTGTGATTTTATAG 1019
Qy 1021 TATGCACATGTCAGAAAAGGCAAGACAAATGGCTCTTGTACTGAATACCTTCGGCAAC 1080
Db 1020 TATGCACATGTCAGAAAAGGCAAGACAAATGGCTCTTGTACTGAATACCTTCGGCAAC 1079
Qy 1081 TTATTGGGCTTCATTTCTGACAGACAGAGATTTGACATCAATATTTGAGAGCTTGGCTA 1140
Db 1080 TTATTGGGCTTCATTTCTGACAGACAGAGATTTGACATCAATATTTGAGAGCTTGGCTA 1139
Qy 1141 GAATGATTTACATGTTAGTGTGATGCACTGTGAGAAATGGTTTTTACTGATCAGCAAT 1200
Db 1140 GAATGATTTACATGTTAGTGTGATGCACTGTGAGAAATGGTTTTTACTGATCAGCAAT 1199
Qy 1201 TCATCTCAGATGAATCTTTTATGCTTTTATGCTAGCATATCTGAATTTACTTTATA 1260
Db 1200 TCATCTCAGATGAATCTTTTATGCTTTTATGCTAGCATATCTGAATTTACTTTATA 1259
Qy 1261 AAGATGGTTTTAGAAAGCTTTGCTAAAAATTTGGCTAGGAATGGTAACCTTCATTTTCA 1320
Db 1260 AAGATGGTTTTAGAAAGCTTTGCTAAAAATTTGGCTAGGAATGGTAACCTTCATTTTCA 1319
Qy 1321 GTTGCAAGGGGTAGAAAATTAATATGCTGTGTTGTTATGTTTATGTTAACAATTTATTAG 1380
Db 1320 GTTGCAAGGGGTAGAAAATTAATATGCTGTGTTGTTATGTTTATGTTAACAATTTATTAG 1379
Qy 1381 GTACTATCATGAATGATTTTAAATATTTTTCATATTTCTGTGACAGCAATTTATAATTG 1440
Db 1380 GTACTATCATGAATGATTTTAAATATTTTTCATATTTCTGTGACAGCAATTTATAATTG 1439
Qy 1441 CAACAAGTGGATGCCATTTAGCCAGTGGGAAGCTTGGAACTCAGGTACCCCTTGAG 1500
Db 1440 CAACAAGTGGATGCCATTTAGCCAGTGGGAAGCTTGGAACTCAGGTACCCCTTGAG 1499
Qy 1501 GATATGCTGCAGCAGCTCTTTTGTGCTTAACTGTTAACTGTAATTTATAGACAGCTAAA 1560
Db 1500 GATATGCTGCAGCAGCTCTTTTGTGCTTAACTGTTAACTGTAATTTATAGACAGCTAAA 1559
Qy 1561 TCCTTAATTTGGATCTGGAATGATTTAGTGTGATGACCTTGTAACAATTTCCAG 1620
Db 1560 TCCTTAATTTGGATCTGGAATGATTTAGTGTGATGACCTTGTAACAATTTCCAG 1619
Qy 1621 GGCATCTGGGTTTGGTCTAGTGTGATTTGAAAACACAGAAACAGAGATCCAGCTGAAAA 1680
Db 1620 GGCATCTGGGTTTGGTCTAGTGTGATTTGAAAACACAGAAACAGAGATCCAGCTGAAAA 1679
Qy 1681 GAGTGATCTCAATATCTCAATCTGATGCTGCTCAACTCAAGCAGAGTTTCTCACTGTG 1740
Db 1680 GAGTGATCTCAATATCTCAATCTGATGCTGCTCAACTCAAGCAGAGTTTCTCACTGTG 1739
Qy 1741 GCATGTCATCATGAACCTAGTAGAGGGATGTTGTTATTTATACAAATTTAATACA 1800
Db 1740 GCATGTCATCATGAACCTAGTAGAGGGATGTTGTTATTTATACAAATTTAATACA 1799
Qy 1801 ATGTCCTTACATTTGATAAAATCTTAAAGAGCAAACTGCAATTTTCTGCAATCCACAT 1860
Db 1800 ATGTCCTTACATTTGATAAAATCTTAAAGAGCAAACTGCAATTTTCTGCAATCCACAT 1859
Qy 1861 TCCATCATATAGAACTAGATATTTATCTATGAGATATTAATGTCGAGAGACTT 1920
Db 1860 TCCATCATATAGAACTAGATATTTATCTATGAGATATTAATGTCGAGAGACTT 1919
Qy 1921 TCATCTCTGATTCGCTGTTGTTCTTAGGGTTCTTAGCACTGTGCTGCACAGCAATGTG 1980
Db 1920 TCATCTCTGATTCGCTGTTGTTCTTAGGGTTCTTAGCACTGTGCTGCACAGCAATGTG 1979

Qy 1981 ATATGTGAATAAATGGATTCCTCTATAGCTAATGAGTTCCCTCTGGGAGAGTTCTG 2040
Db 1980 ATATGTGAATAAATGGATTCCTCTATAGCTAATGAGTTCCCTCTGGGAGAGTTCTG 2039
Qy 2041 GTACTCAATCACAATGCCAGATGGTTTATAGGGCTATTTGTGTAAGTAAGTGGTAAGA 2100
Db 2040 GTACTCAATCACAATGCCAGATGGTTTATAGGGCTATTTGTGTAAGTAAGTGGTAAGA 2099
Qy 2101 TGCATGGAAGTGAAGTGTGTTTCTTTCATCTTATGGAACCTCTGATGATGCTGCTTTTG 2160
Db 2100 TGCATGGAAGTGAAGTGTGTTTCTTTCATCTTATGGAACCTCTGATGATGCTGCTTTTG 2159
Qy 2161 TATGGAATAAATTTTGGTGCATATGATGTCATTCATTCATTCATTCATTCATTCATTCAT 2220
Db 2160 TATGGAATAAATTTTGGTGCATATGATGTCATTCATTCATTCATTCATTCATTCATTCAT 2219
Qy 2221 TTGTATTTATGATATATATATACCTGTCACGCTCTAGTTCATTCATTCATTCATTCAT 2280
Db 2220 TTGTATTTATGATATATATATACCTGTCACGCTCTAGTTCATTCATTCATTCATTCAT 2279
Qy 2281 TTTTGTACATATTTTACTTGAATAATTTTAAATGGAAATTTTAAATPAAACATTTGATAGT 2340
Db 2280 TTTTGTACATATTTTACTTGAATAATTTTAAATGGAAATTTTAAATPAAACATTTGATAGT 2339

Qy 2341 TTACAT 2346
Db 2340 TTACAT 2345

RESULT 3

AAS94792

ID AAS94792 standard; DNA; 2321 BP.

XX AAS94792;

XX 14-FEB-2002 (first entry)

XX Human DNA sequence #47 expressed during foam cell differentiation.

XX Human; foam cell differentiation; atherosclerosis; cerebral stroke;
XX cardiovascular disorder; coronary artery disease; gene therapy; ds.

XX Homo sapiens.

XX WO200177389-A2.

XX 18-OCT-2001.

XX 04-APR-2001; 2001WO-US011128.

XX 05-APR-2000; 2000US-0195106P.

XX (INCY-) INCYTE GENOMICS INC.

XX Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;
PI Tai J;

XX WPI; 2002-010925/01.

XX Composition useful for diagnosis of conditions, disorders or diseases
XX associated with atherosclerosis, comprises several polynucleotides that
XX are differentially expressed in foam cell development.

XX Claim 1; Page 102-103; 315pp; English.

XX The present invention relates to the isolation of human polynucleotide
XX sequences that are differentially expressed during foam cell
XX differentiation. The polynucleotide sequences of the invention or a
XX composition comprising these polynucleotides are useful as a high
XX throughput method for detecting altered expression of one or more
XX polynucleotides in a sample. The polynucleotides can be used in the
XX diagnosis of disorders associated with foam cell development such as

CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as
CC coronary artery disease. The polynucleotide sequences can also be used as
CC PCR primers and probes. The polynucleotides of the invention are also
CC useful in gene therapy. AAS94746-AAS95021 represent the human
CC polynucleotide sequences of the invention which are differentially
CC expressed during foam cell differentiation
XX

XX .SQ Sequence 2321 BP; 700 A; 358 C; 421 G; 841 T; 0 U; 1 Other;

Query Match 98.2%; Score 2304.4; DB 6; Length 2321;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2316; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 28 CAGGTATCTCTTTTGCAGTGGTCTGATTTTGGTACCTTGTGACAGGGTATTAT 87
DB 4 CAGGTATCTCTTTTGCAGTGGTCTGATTTTGGTACCTTGTGACAGGGTATTAT 62
QY 88 TACAGCATCTTGGGAAACCTATTAGCCCTTGCATGTTTAAAGCTGATTAATTTG 147
DB 63 TACAGCATCTTGGGAAACCTATTAGCCCTTGCATGTTTAAAGCTGATTAATTTG 122
QY 148 GGTGTGAGTGGTCTGACCTTAAATGTATTAATAAATTTAGACATCAAAATTTCTTACT 207
DB 123 GGTGTGAGTGGTCTGACCTTAAATGTATTAATAAATTTAGACATCAAAATTTCTTACT 182
QY 208 AACTAATTTTATAGATGATCTTGGAGCACAGTCAATACACCTGGGAGGCAATGCA 267
DB 183 AACTAATTTTATAGATGATCTTGGAGCACAGTCAATACACCTGGGAGGCAATGCA 242
QY 268 ATGTGTTACCTGGTCTAGGTTGAACCTGCTTATTTCAAAAGATTTCTGAATTAATTT 327
DB 243 ATGTGTTACCTGGTCTAGGTTGAACCTGCTTATTTCAAAAGATTTCTGAATTAATTT 302
QY 328 TTCCCTAGAAATTTCTTCAATCCAAAGTCAAAACATCTTGAAGATGAAGACAGATT 387
DB 303 TTCCCTAGAAATTTCTTCAATCCAAAGTCAAAACATCTTGAAGATGAAGACAGATT 362
QY 388 GTTCCCAATGATGTATGCTCATACCTGACATGAAACGATCTATGTTAAATGACTGTGAT 447
DB 363 GTTCCCAATGATGTATGCTCATACCTGACATGAAACGATCTATGTTAAATGACTGTGAT 422
QY 448 ATGAATATTTCAAGTACTACCCCAATTAATCTTCTTATGCTCTGAAGAGGAAGCA 507
DB 423 ATGAATATTTCAAGTACTACCCCAATTAATCTTCTTATGCTCTGAAGAGGAAGCA 482
QY 508 ATGTAATCACTATGATTTATGCAAAACCAAGCAATTTCTCAACCAATTTTAAAGTAATC 567
DB 483 ATGTAATCACTATGATTTATGCAAAACCAAGCAATTTCTCAACCAATTTTAAAGTAATC 542
QY 568 TGATCTCTCTTGGAGAAATTTGATCTAATAGTCTTCTTATGATGATTTTAACTA 627
DB 543 TGATCTCTCTTGGAGAAATTTGATCTAATAGTCTTCTTATGATGATTTTAACTA 602
QY 628 CTGGTATAAATCAAAATTTCTAATAATTTCTAATAGTCTTAAAGAACTGGGTTCTTCT 687
DB 603 CTGGTATAAATCAAAATTTCTAATAATTTCTAATAGTCTTAAAGAACTGGGTTCTTCT 662
QY 688 TTGATGTTATTCATGTTCAAGAGGAACCAACATTTTACTCTTTTGAAGCAATTCCTAGA 747
DB 663 TTGATGTTATTCATGTTCAAGAGGAACCAACATTTTACTCTTTTGAAGCAATTCCTAGA 722
QY 748 ATCTATAGTATGATCAGGATATATTTGCTTTAAATAATATTTTGGTTATTTGAATACA 807
DB 723 ATCTATAGTATGATCAGGATATATTTGCTTTAAATAATATTTTGGTTATTTGAATACA 782
QY 808 GACATTTGGCTCCAAATTTTCTCTTGGACATAGTATGATCTTCTCAGTAACTCTCA 867
DB 783 GACATTTGGCTCCAAATTTTCTCTTGGACATAGTATGATCTTCTCAGTAACTCTCA 842
QY 868 ACATTTGGAACTTTGCAAAATGAGCATCAATGTTTAAAGGCTGATCAATTTAATGCT 927
DB 843 ACATTTGGAACTTTGCAAAATGAGCATCAATGTTTAAAGGCTGATCAATTTAATGCT 902

QY 928 ATGAGATACATTTGTTTCTCCTATGCAAAACAGGTGAACAAACAGTAGTTGTTTTTACT 987
DB 903 ATGAGATACATTTGTTTCTCCTATGCAAAACAGGTGAACAAACAGTAGTTGTTTTTACT 962
QY 988 GATATCTAAATGTTGGCTACCTGTGATTTTATAGTATGCACATGTGAGAAAAGCAAGAC 1047
DB 963 GATATCTAAATGTTGGCTACCTGTGATTTTATAGTATGCACATGTGAGAAAAGCAAGAC 1022
QY 1048 AATAGGCTCTGTTGTTACTGATCTTGGCAAACTTATGGTCTTCAATTTCTGACAGAC 1107
DB 1023 AATAGGCTCTTGTACTGATCTTGGCAAACTTATGGTCTTCAATTTCTGACAGAC 1082
QY 1108 AGGATTTGACTCAATATTTGTAGAGCTTGGCTAGAGTGAATTAATGATGATGATGACT 1167
DB 1083 AGGATTTGACTCAATATTTGTAGAGCTTGGCTAGAGTGAATTAATGATGATGACT 1142
QY 1168 GGTAGAAATGTTTTAGTTATGACTCAGAAATTCATCTCAGGATGAATCTTTTATGCT 1227
DB 1143 GGTAGAAATGTTTTAGTTATGACTCAGAAATTCATCTCAGGATGAATCTTTTATGCT 1202
QY 1228 TTTTATTTGAAGCATATCTGAATTTTACTTTTATAAGATGGTTTTAGAAAGCTTTTGTCTAA 1287
DB 1203 TTTTATTTGAAGCATATCTGAATTTTACTTTTATAAGATGGTTTTAGAAAGCTTTTGTCTAA 1262
QY 1288 AATTTGGCTTAGGAATGTTTCAATTTTCACTTGGCAAGGGGTAGAAAATAATATG 1347
DB 1263 AATTTGGCTTAGGAATGTTTCAATTTTCACTTGGCAAGGGGTAGAAAATAATATG 1322
QY 1348 TGTGTTGTTATGTTTATGTTAAACATATTTATTTAGGTACTATCTATGAATGTTTAAATAT 1407
DB 1323 TGTGTTGTTATGTTTATGTTAAACATATTTATTTAGGTACTATCTATGAATGTTTAAATAT 1382
QY 1408 TTTTTCATATCTGTGACAGCATTTTATAATTTGCAACAGGTGGAGTCCATTTAGCCAGT 1467
DB 1383 TTTTTCATATCTGTGACAGCATTTTATAATTTGCAACAGGTGGAGTCCATTTAGCCAGT 1442
QY 1468 GGGAAAGCTTTGGAACTCAGGTTTACCTTTGAAGGATATGCTGGCAGCCATCTCTTTGATC 1527
DB 1443 GGGAAAGCTTTGGAACTCAGGTTTACCTTTGAAGGATATGCTGGCAGCCATCTCTTTGATC 1502
QY 1528 TGTGCTTAACTGTTAATTTATAGACCAAGTAAATCCCTTAATCTTGGATCGGAATGATTA 1587
DB 1503 TGTGCTTAACTGTTAATTTATAGACCAAGTAAATCCCTTAATCTTGGATCGGAATGATTA 1562
QY 1588 GTTATGACCTTTGTACCAATTTCCAGAAATTTGAGGGCACTGGTGGTTGGTCTAGTATG 1647
DB 1563 GTTATGACCTTTGTACCAATTTCCAGAAATTTGAGGGCACTGGTGGTTGGTCTAGTATG 1622
QY 1648 AATACACAGAACAGAGATCCAGCTGAAAGAGTATCCCTCAATATCTTAACTAACT 1707
DB 1623 AATACACAGAACAGAGATCCAGCTGAAAGAGTATCCCTCAATATCTTAACTAACT 1682
QY 1708 GGTCTCAACTCAAGCAGAGTTTCTTCACTTGGCACTGTGATCATGAAACTTAGTAGAG 1767
DB 1683 GGTCTCAACTCAAGCAGAGTTTCTTCACTTGGCACTGTGATCATGAAACTTAGTAGAG 1742
QY 1768 GGGATTTGCTGTTTATTAATAAATTTAATACATGTTTACATTTGATAAATTTCTTAAA 1827
DB 1743 GGGATTTGCTGTTTATTAATAAATTTAATACATGTTTACATTTGATAAATTTCTTAAA 1802
QY 1828 GAGCAAACTGCAATTTTATTTCTGCATCCCATTTCCCAATCATATTTAGAACTAAGATTTT 1887
DB 1803 GAGCAAACTGCAATTTTATTTCTGCATCCCATTTCCCAATCATATTTAGAACTAAGATTTT 1862
QY 1888 ATCTATGAGGATATAAATTTGGTGCAGAGAGACTTTTCACTGTGGATGCGTTGTTCTTAG 1947
DB 1863 ATCTATGAGGATATAAATTTGGTGCAGAGAGACTTTTCACTGTGGATGCGTTGTTCTTAG 1922
QY 1948 GGTTCCTAGCACTGATGCTGCACAGCATGTGATATGTGAAAATAAATCGAATCTTCTTA 2007
DB 1923 GGTTCCTAGCACTGATGCTGCACAGCATGTGATATGTGAAAATAAATCGAATCTTCTTA 1982
QY 2008 TAGCTAAATAGTTCCCTCTCTGGGAGAGATTCTGCTACTGCAATCAACATGCCAGATGGTG 2067

[illegible]

Dd		1500 TATTATACCTGTGTCAGCGTCTTAGTTGGCTCAACCATTATAAACCATTTTTTGTTACATATT 1567
Qy		2294 TTACTTGAAAATAATTTAAATGGAATTTAAAAATAAACAATTTGCATAGTTTACAT 2346 1560 TTACTTGAAAATAATTTAAATGGAATTTAAATGGAATTTAAATAAACAATTTGCATAGTTTACAT 1612
Db		
RESULT 6		
AAD29955		
ID ID		AAD29955 standard; DNA; 1626 BP.
XX AC		AAD29955;
XX DT		17-MAY-2002 (first entry)
XX DE		Human pharmaceutical compound #24 for cancer treatment.
XX XX		Human; pharmaceutical composition; compound I; tumour; psoriasis; cancer; rheumatoid arthritis; vascular endothelial growth factor; VEGF; therapy;
KW KW		neovascular glaucoma; compound II; angiotensin/Tie receptor system;
KW KW		retinopathy; glomerulonephritis; diabetic nephropathy; nephrosclerosis;
KW KW		chronic microangiopathic syndrome; transplantation; glomerulopathy;
KW KW		fibrotic disease; cirrhotic liver; proliferative disease; nephrotropic;
KW KW		ophthalmological; arteriosclerosis; cystostatic; hepatotropic; oedema; ds.
OS OS		Homo sapiens.
XX XX		WO2001197850-A2.
PX PX		27-DEC-2001.
PD PD		20-JUN-2001; 2001WO-BP006976.
PF PF		23-JUN-2000; 2000EP-00250194.
PR PR		28-JUN-2000; 2000EP-00250214.
PX PX		(SCHD) SCHERING AG. (SIEM/) SIEMEISTER G. (HABE/) HABEREY M. (THIE/) THIERRACH K.
PA PA		Siemeister G, Haberey M, Thierauch K; WPI; 2002-179543/23.
XI XI		Novel composition useful for treating cancer, comprises agents interfering with vascular endothelial growth factor/VEGF receptor system activity and agents interfering with Angipoietin/Tie receptor system function.
XX XX		Claim 12; Page 63; 79pp; English.
PS PS		The present invention relates to a pharmaceutical composition comprising a combination of substances (compound I) interfering with the biological activity of vascular endothelial growth factor (VEGF)/VEGF receptor systems and substances (compound II) interfering with the biological function of angiopoietin/Tie receptor systems. The pharmaceutical composition is useful for the production of a medicament for the treatment of tumours, cancers, psoriasis, arthritis, such as rheumatoid arthritis, haemangoma, angiofibroma, eye disease such as retinitis, neovascular glaucoma, kidney disease such as chronic renal insufficiency, malignant nephrosclerosis, thrombotic microangiopathic syndrome, transplant rejection and glomerulopathies, fibrotic diseases such as cirrhotic liver, mesangial cell proliferation diseases, arteriosclerosis, damage of nerve tissues, suppression of the ascites formation in patients and suppression of VEGF oedemas. The present sequence is human pharmaceutical compound used in the invention
CC CC		Sequence 1626 BP; 493 A; 239 C; 306 G; 588 T; 0 U; 0 Other;
XX XX		Query Match 68.2%; Score 1599.4; DB 6; Length 1626; Best Local Similarity 99.94; Pred. No. 0;

Query Match	68.2%;	Score 1599.4;	DB 6;	Length 1626;
Best Local Similarity	99.9%;	Pred. No. 0;		

Qy	1814	ATATAATCTTAAAGACAAAAC	CTGCATTTTATTTCTGCATCCCAATCCCAATCATATTA	1873
Db	1080	ATATAATCTTAAAGACAAAAC	CTGCATTTTATTTCTGCATCCCAATCCCAATCATATTA	1139
Qy	1874	GAACCTAAGATATTTATCTAT	GAAGATATAAAATGGTGCAGAGAGACTTTTCATCTGTGGATT	1933
Db	1140	GAACCTAAGATATTTATCTAT	GAAGATATAAAATGGTGCAGAGAGACTTTTCATCTGTGGATT	1199
Qy	1934	GGCTGTGTTCTTAGGGTTCT	AGCACTGATGCGCTGCACAAGCATGTGATATGTCAATAA	1993
Db	1200	GGCTGTGTTCTTAGGGTTCT	AGCACTGATGCGCTGCACAAGCATGTGATATGTCAATAA	1259
Qy	1994	AATGATTCCTTATAGCTAAAT	AGATCCCTCTCGGGAGAGTTCTTGGTACTGCAATCAC	2053
Db	1260	AATGATTCCTTATAGCTAAAT	AGATCCCTCTCGGGAGAGTTCTTGGTACTGCAATCAC	1319
Qy	2054	AATGCCAGATGGTCTTATAT	GGGCTATTTCTGTGTAAGTAAAGTGGTAAGATGCTTATGAAGTAA	2113
Db	1320	AATGCCAGATGGTCTTATAT	GGGCTATTTCTGTGTAAGTAAAGTGGTAAGATGCTTATGAAGTAA	1379
Qy	2114	GTGTTGTTGTTTTCATCTTA	TGAAATCTCTTGATGCAATGCTGCTTTTGTATGGAATAAATT	2173
Db	1380	GTGTTGTTGTTTTCATCTTA	TGAAATCTCTTGATGCAATGCTGCTTTTGTATGGAATAAATT	1439
Qy	2174	TTGTGCAATATGATGTCATT	CAACTTTTGCATTTGAATTTTGGTTGTATTTATATG	2233
Db	1440	TTGTGCAATATGATGTCATT	CAACTTTTGCATTTGAATTTTGGTTGTATTTATATG	1499
Qy	2234	TATATACCTCTCAGCTTCT	AGTTTCTGCTTCAACCACTTTTATAACCATTTTGTACATATT	2293
Db	1500	TATATACCTCTCAGCTTCT	AGTTTCTGCTTCAACCACTTTTATAACCATTTTGTACATATT	1559
Qy	2294	TTACTTGAAAATATTTTAAAT	TGAAAATTTTAAATAAACAATTTGATAGTTTACAT	2346
Db	1560	TTACTTGAAAATATTTTAAAT	TGAAAATTTTAAATAAACAATTTGATAGTTTACAT	1612
RESULT 7				
AAH34083/C				
ID: AAH34083 standard; cDNA; 665 BP.				
XX	AAH34083;			
XX	AC			
XX	DT			
XX	03-SEP-2001 (first entry)			
XX	Human colon cancer antigen encoding cDNA SEQ ID NO:1165.			
DE	Human; colon cancer antigen; diagnosis; detection;			
XX	colorectal carcinoma; ss.			
XX	Homo sapiens.			
OS	WO200122920-A2.			
XX	PN			
XX	PD			
XX	05-APR-2001.			
XX	28-SEP-2000; 2000WO-US026524.			
PF	XX			
XX	29-SEP-1999; 99US-0157137P.			
PR	XX			
XX	03-NOV-1999; 99US-0163280P.			
XX	XX			
XX	(HUMA-) HUMAN GENOME SCI INC.			
PA	XX			
XX	Ruben SM, Barash SC, Birse CE, Rosen CA;			
PI	XX			
XX	WPI; 2001-235357/24.			
DR	XX			
XX	P-PSDB; AAG74678.			
XX	Nucleic acids encoding 4277 human colon cancer-associated polypeptides,			
PT	XX			
XX	useful for preventing, diagnosing and/or treating colorectal cancers.			
PS	Claim 1; Page 2972; 9803pp; English.			

RESULT 8
ABZ08185

QY 708 AAGGAAACACACCTTCTCTTTAGGACAACTCTAGATCTATAGTATCAGAT 767
 Db 241 AAGGAAACACACCTTCTCTTTAGGACAACTCTAGATCTATAGTATCAGAT 300
 QY 768 ATATTTGGCTTAAATATATTTGGTTATTTGGAATACAGACATTTGGCTCCAAATTTTC 827
 Db 301 ATATTTGGCTTAAATATATTTGGTTATTTGGAATACAGACATTTGGCTCCAAATTTTC 360
 QY 828 ATCTTTGACAAATAGTATGACCTTTCTAGAACCTTCTCAACATTTGGGAACCTTTGCAAA 887
 Db 361 ATCTTTGACAAATAGTATGACCTTTCTAGAACCTTCTCAACATTTGGGAACCTTTGCAAA 420
 QY 888 TAGAGCATCATATGTGTATAGGCTGTATCATTTAATGCTATGAGATACATTTGTTTCTC 947
 Db 421 TAGAGCATCATATGTGTATAGGCTGTATCATTTAATGCTATGAGATACATTTGTTTCTC 480
 QY 948 CCTATGCCAAACAGGTCAGAAACCTAGTGTGTTTTTACTGATACTAAATGTTGGCTACC 1007
 Db 481 CCTATGCCAAACAGGTCAGAAACCTAGTGTGTTTTTACTGATACTAAATGTTGGCTACC 540
 QY 1008 TGTGATTTTATAGTATGACATGTGCAGAAAGGCAAGACAAATGGCTCTTG 1060
 Db 541 TGTGATTTTATAGTATGACATGTGCAGAAAGGCAAGACAAATGGCTCTTG 593

RESULT 9

AAL22443
 ID AAL22443 standard; cDNA; 608 BP.
 XX
 AC AAL22443;

XX 07-DEC-2001 (first entry)
 XX
 DE Human breast cancer expressed polynucleotide 14900.
 XX
 KM Human; breast cancer; cell marker; cytostatic; ss.
 XX
 OS Homo sapiens.

XX
 PN WO200151628-A2.
 XX
 PD 19-JUL-2001.

XX 10-JAN-2001; 2001WO-US000798.
 XX
 PF 14-JAN-2000; 2000US-0176077P.
 XX
 PR 14-MAR-2000; 2000US-0189167P.
 XX
 PR 24-MAR-2000; 2000US-0192099P.
 XX
 PR 29-MAR-2000; 2000US-0193480P.
 XX
 PR 15-MAY-2000; 2000US-0205230P.
 XX
 PR 09-JUN-2000; 2000US-0211315P.
 XX
 PR 25-JUL-2000; 2000US-0220534P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 XX Lillie J. Xu Y, Wang Y, Steinmann K;
 XX
 XX WPI; 2001-451856/48..

XX New peptide useful as a marker for the diagnosis of breast cancer.
 XX
 XX Claim 1; Page 2691; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides
 CC (AAL07544-AAL26789) and methods of assessing whether a patient is
 CC afflicted with breast cancer by examining the correlation between the
 CC expression of certain markers and the cancerous state of breast cells.
 CC The polynucleotides and encoded polypeptides are potential markers for
 CC detecting, diagnosing, and monitoring, characterising treating and
 CC potentially preventing breast cancer. The polynucleotides and encoded
 CC polypeptides are also useful for isolating compounds with cytostatic
 CC activity

XX SQ Sequence 608 BP; 198 A; 101 C; 87 G; 222 T; 0 U; 0 Other;
 Query Match 24.9%; Score 584.8; DB 4; Length 608;
 Best Local Similarity 98.8%; Pred. No. 3.4e-114;
 Matches 589; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 468 CCCCAATAAATCTTCTTATTTGCTCTGAAAGAAAGAAAGCAATGTAATCACTATGATTAT 527
 Db 10 CCCAAATAAATCTTCTTATTTGCTCTGAAAGAAAGAAAGCAATGTAATCACTATGATTAT 69
 QY 528 TGCACAAACACAGAAATCTCCAAATTTTAAAGTAATCTGATCCTCTCTTCTGGAGAAA 587
 Db 70 TGCACAAACACAGAAATCTCCAAATTTTAAAGTAATCTGATCCTCTCTTCTGGAGAAA 129
 QY 588 ATTTGTAACCTAATAGTTTTTCTCTTATGAATGTTATTTACTACTGCTATATAATCAAAATTTCT 647
 Db 130 ATTTGTAACCTAATAGTTTTTCTCTTATGAATGTTATTTACTACTGCTATATAATCAAAATTTCT 189
 QY 648 ATAAATTTCTTACTTAAAGTCTTAAAGAACTGGGTCTTCTCTTATGATGTTATTCATGTTTCAG 707
 Db 150 ATAAATTTCTTACTTAAAGTCTTAAAGAACTGGGTCTTCTCTTATGATGTTATTCATGTTTCAG 249
 QY 708 AAAGGAAACAAACACTTTTACTCTTTTAGGACAAATTTCTAGAAATCTATAGTAGTATCAGGAT 767
 Db 250 AAAGGAAACAAACACTTTTACTCTTTTAGGACAAATTTCTAGAAATCTATAGTAGTATCAGGAT 309
 QY 768 ATATTTTCTTAAATATATTTTGGTATTTTGAATACAGACATTTGGCTCCAAATTTTC 827
 Db 310 ATATTTTCTTAAATATATTTTGGTATTTTGAATACAGACATTTGGCTCCAAATTTTC 369
 QY 828 ATCTTTGCACAATAGTATGACTTTTCTACTAGAACTTCTCAACATTTGGGAACCTTTGCAAA 887
 Db 370 ATCTTTGCACAATAGTATGACTTTTCTACTAGAACTTCTCAACATTTGGGAACCTTTGCAAA 429
 QY 888 TATGAGCATCATATGTGTATAGGCTGTATCATTTAATGCTATGAGATACATTTGTTTCTC 947
 Db 430 TATGAGCATCATATGTGTATAGGCTGTATCATTTAATGCTATGAGATACATTTGTTTCTC 489
 QY 948 CCTATGCCAAACAGGTCAGAAACCTAGTGTGTTTTTACTGATACTAAATGTTGGCTACC 1007
 Db 490 CCTATGCCAAACAGGTCAGAAACCTAGTGTGTTTTTACTGATACTAAATGTTGGCTACC 549
 QY 1008 TGTGATTTTATAGTATGACATGTGCAGAAAGGCAAGACAAATGGCTCTCTGAC 1063
 Db 550 TGTGATTTTATAGTATGACATGTGCAGAAAGGCAAGACAAATGGCTCTCTGAC 605

RESULT 10

ABK27756/c
 ID ABK27756 standard; cDNA; 586 BP.
 XX
 AC ABK27756;

XX 09-APR-2002 (first entry)
 XX
 DE Human colon cancer expressed sequence tag, Seq ID no 193.
 XX
 KW Human; colon cancer; T cell expansion; tumour; EST; gene; ss;
 XX expressed sequence tag.
 XX
 OS Homo sapiens.

XX WO200156390-A2.
 XX
 PD 20-DEC-2001.

XX 08-JUN-2001; 2001WO-US018577.
 XX
 PR 09-JUN-2000; 2000US-0210821P.
 XX
 PR 18-DEC-2000; 2000US-0256571P.
 XX
 PR 10-MAY-2001; 2001US-0250240P.

PA	(CORI-) CORIXA CORP.	
XX	Jiang Y, Hepler WT, Clapper JD, Wang A, Secrist H;	
PI	WPI; 2002-139708/18.	
XX	Novel isolated polynucleotide encoding a polypeptide comprising a portion	
XX	of colon tumor protein, useful for detection, diagnosis and therapy of	
PT	human colon cancer.	
PT	Claim 1; Page 202; 220pp; English.	
PS	The invention relates to an isolated polynucleotide (I) encoding a	
XX	polypeptide (II) comprising at least a portion of a colon tumour protein.	
XX	(I), (II) and antibody (III) to (I) are useful for determining the	
CC	presence of a cancer in a patient. (I), (II) or antigen presenting cells	
CC	expressing (I) is useful for stimulating and/or expanding T cells	
CC	specific for a tumour protein, by contacting T cells with (I), (II) or	
CC	antigen-presenting cells that express (I), under conditions and for a	
CC	time sufficient to permit the stimulation and/or expansion of T cells.	
CC	(I), (II) or antigen presenting cells that express (II) are useful for	
CC	treating colon cancer in a patient by incubating CD4+ and/or CD8+ T cells	
CC	isolated from a patient with (I), (II) or antigen presenting cells that	
CC	express (II), such that T cells proliferate, and administering to the	
CC	patient an effective amount of the proliferated T cells, thus inhibiting	
CC	the development of a cancer in the patient. (I) or (II) is useful in	
CC	vaccines and pharmaceutical compositions for prevention and treatment of	
CC	colon malignancies and for the diagnosis and monitoring of such cancers.	
CC	(I), (II) or (III) is useful for detection, diagnosis and/or therapy of	
CC	human colon cancer. (I) is useful as a probe or primer for nucleic acid	
CC	hybridisation, and in the design and preparation of ribozyme molecules	
CC	for inhibiting expression of (II) in tumour cells. ABK27564-ABK27807	
CC	represent novel human colon cancer coding sequences and primers of the	
CC	invention	
XX	Sequence 586 BP; 214 A; 88 C; 97 G; 187 T; 0 U; 0 Other;	
SQ	Query Match 24.8%; Score 582.8; DB 6; Length 586;	
	Best Local Similarity 99.7%; Pred. No. 8.9e-114;	
	Matches 584; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	476 AACTTCTTATTCCTCTGAAGAAGAAAGCAATGTAATCACTATGATTATTCACAAA 535	
DB	586 AACTTCTTATTCCTCTGAAGAAGAAAGCAATGTAATCACTATGATTATTCACAAA 527	
QY	536 CAACCAAGATCTCCAAACAAATTTAAGTAATCTGATCCTCTTCTGGAGAAATTTGTTAC 595.	
DB	526 CAACCAAGATCTCCAAACAAATTTAAGTAATCTGATCCTCTTCTGGAGAAATTTGTTAC 467	
QY	596 CTAATAGTTTTTCTTATGATGTTATTTACTACTGTTATTAATCAAAATTTCTATAAATTT 655	
DB	466 CTAATAGTTTTTCTTATGATGTTATTTACTACTGTTATTAATCAAAATTTCTATAAATTT 407	
QY	656 CTTACTTAAGTCTTAAGAACTGGGTTCTTCTTTGATGTTATTCATGTTTCAGAAAGGAAA 715	
DB	406 CTTACTTAAGTCTTAAGAACTGGGTTCTTCTTTGATGTTATTCATGTTTCAGAAAGGAAA 347	
QY	716 CAACACTTTACTTTTAGACAAATCTAGATCTATAGTATGATGATGATGATGATGATGATG 775	
DB	346 CAACACTTTACTTTTAGACAAATCTAGATCTATAGTATGATGATGATGATGATGATGATG 287	
QY	776 CTTTAAATATATTTTGGTTATTTTGAATACAGACATTTGGCTCCAAATTTTCTATCTTTCG 835	
DB	286 CTTTAAATATATTTTGGTTATTTTGAATACAGACATTTGGCTCCAAATTTTCTATCTTTCG 227	
QY	836 ACAATAGTATGATCTTTTCACTAGAACTTCTCAACATTTGGAACTTTGCAATATGAGCA 895	
DB	226 ACAATAGTATGATCTTTTCACTAGAACTTCTCAACATTTGGAACTTTGCAATATGAGCA 167	
QY	896 TCATATGCTGTAAGGCTGTATCATTTAATGCTATGAGATACATTTGTTTCTCCCTATGCC 955	
DB	166 TCATATGCTGTAAGGCTGTATCATTTAATGCTATGAGATACATTTGTTTCTCCCTATGCC 107	
QY	956 AAACAGGTGAACAAACGCTAGTCTTTTCTTACTGATCTAAATGTTGGCTACCTGCTGATTT 1015	
DB	106 AAACAGGTGAACAAACGCTAGTCTTTTCTTACTGATCTAAATGTTGGCTACCTGCTGATTT 47	
QY	1016 TATAGTATGACATGTCAGAAAAAGGCAAGACAAATGGCTCTTTGT 1061	
DB	46 TATAGTATGACATGTCAGAAAAAGGCAAGACAAATGGCTCTTTGT 1	
RESULT 11		
ABK39231		
ID	ABK39231 standard; cDNA; 675 BP.	
XX	ABK39231;	
AC	ABK39231;	
XX	21-MAY-2002 (first entry)	
DT	DNA encoding lung tumour protein associated contig 89.	
DE	Lung tumour; cancer; T cell; immune response stimulator; cytostatic;	
XX	contig; ss.	
KW	Homo sapiens.	
XX	WO200204514-A2.	
XX	17-JAN-2002.	
PD	10-JUL-2001; 2001WO-US022058.	
PP	11-JUL-2000; 2000US-00614124.	
XX	29-AUG-2000; 2000US-00651563.	
PR	08-SEP-2000; 2000US-00658824.	
PR	26-SEP-2000; 2000US-00671325.	
PR	06-OCT-2000; 2000US-00677419.	
PR	30-OCT-2000; 2000US-00702705.	
PR	13-DEC-2000; 2000US-00736457.	
PR	03-MAY-2001; 2001US-00849626.	
XX	(CORI-) CORIXA CORP.	
PA	Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;	
PI	Markezakis M, Carter D, Fanger GR, Vedwick TS, Bangur CS, McNabb A;	
PI	Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;	
XX	WPI; 2002-164634/21.	
DR	Novel polynucleotide encoding a lung tumor polypeptide useful for	
PT	stimulating and/or expanding T cells specific for a tumor protein.	
PT	Claim 1; SEQ ID NO 1269; 223pp; English.	
PS	The invention describes an isolated polynucleotide and polypeptide useful	
XX	for stimulating and/or expanding T cells specific for a tumour protein	
CC	for determining the presence of a cancer in a patient. A composition	
CC	containing the polynucleotide and/or polypeptide is useful for treating a	
CC	lung cancer in a patient. The polypeptide is useful for removing tumour	
CC	cells from a biological sample. The polynucleotide is also useful as	
CC	probe or primer to detect the level of mRNA encoding a tumour protein.	
CC	This sequence represents a contig of a DNA sequence encoding a lung	
CC	tumour associated protein, described in the method of the invention.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ftp.wipo.int/pub/published_pct_sequences	
XX	Sequence 675 BP; 193 A; 110 C; 133 G; 236 T; 0 U; 3 Other;	
SQ	Query Match 24.8%; Score 581.4; DB 6; Length 675;	
	Best Local Similarity 97.3%; Pred. No. 1.8e-113;	
	Matches 622; Conservative 0; Mismatches 13; Indels 4; Gaps 3;	
QY	1673 CTGAAAAAGAGTGTCTCTCAATATCTCAATATCTCTCACTCAAGCAGATTTCT 1732	

XX Human colon tumour antigen polynucleotide SEQ ID NO:1423.
DE Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
XX Human; colon tumour metastatic antigen; diagnosis; gene; ss.
KW Human; colon tumour metastatic antigen; diagnosis; gene; ss.
XX Homo sapiens.

OS Homo sapiens.

PN WO200196388-A2.

XX 20-DEC-2001.

XX 08-JUN-2001; 2001WO-US018557.

XX 09-JUN-2000; 2000US-0210899P.

XX 20-FEB-2001; 2001US-0270216P.

XX (CORI-) CORIXA CORP.

XX Jiang Y, Harlocker SL, Secret H;

XX WPI; 2002-114514/15.

XX Novel isolated colon tumor polynucleotide differentially expressed in
PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
PT useful for inhibiting development of cancer in patient.

XX Claim 1; SEQ ID NO 1423; 105pp; English.

XX ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
CC which were isolated from human colon tumour and colon metastatic tumour
CC cDNA libraries. (I) have cytostatic activity and can be used in vaccine
CC production. (I) can be used for stimulating and/or expanding T cells
CC specific for a tumour protein on contact with the T cells. They are also
CC useful for inhibiting the development of cancer in a patient. (I) can be
CC used as probes or primers for nucleic acid hybridisation, for preparing
CC mutant species primers, or primers for use in genetic constructions. (I)
CC can be used in the diagnosis of a colon tumour

XX Sequence 586 BP; 189 A; 98 C; 85 G; 211 T; 0 U; 3 Other;

XX Query Match 24.2%; Score 588; DB 6; Length 586;

XX Best Local Similarity 99.0%; Pred. No. 1.2e-110;

XX Matches 580; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 464 ACTACCCCAATAAATCTTCTTATGCTCTGGAAGAGAAAGCAATGTAATCACTATGA 523

DB 1 ACTACCCCAATAAATCTTCTTATGCTCTGGAAGAGAAAGCAATGTAATCACTATGA 60

QY 524 TTATTGCAACAACACAGAAATCTCCAAACAATTTTAAAGTAATCTGATCTCTTCTTGA 583

DB 61 TTATTGCGCAACAACACAGAAATCTCCAAACAATTTTAAAGTAATCTGATCTCTTCTTGA 120

QY 584 GAAATGTTACTTAATAGTCTTTTCTTATGATGTTTACTTCTGATATAAATCAAT 643

DB 121 GAAATGTTACTTAATAGTCTTTTCTTATGATGTTTACTTCTGATATAAATCAAT 180

QY 644 TTCTATAAATTTCTTACTTAATGTTTAAAGCTGGTCTTCTTCTTCTTCTTCTTCTTCT 703

DB 181 TTCTATAAATTTCTTACTTAATGTTTAAAGCTGGTCTTCTTCTTCTTCTTCTTCTTCT 240

QY 704 TCAGAAGGAACAACACATTTTCTTTTAGGACAAATCTTAGAATCTATAGTATCA 763

DB 241 TCAGAAGGAACAACACATTTTCTTTTAGGACAAATCTTAGAATCTATAGTATCA 300

QY 764 GGATATATTTTGGTTTAAATATATTTTGGTTATTTTGAATACAGACATTTGGCTCCAAT 823

DB 301 GGATATATTTTGGTTTAAATATATTTTGGTTATTTTGAATACAGACATTTGGCTCCAAT 360

QY 824 TTTCATCTTTGCAATAGTATGATCTTTTCACTAGACTTCTTCAACATTTGGGACTTTG 883

DB 361 TTTCATCTTTGCAATAGTATGATCTTTTCACTAGACTTCTTCAACATTTGGGACTTTG 420

QY 884 CAAATATGAGCATCATATGTTTAAGGCTG-TATCATTTTAATGCTATGAGATACATTGT 942

DB 421 CAAATATGAGCATCATATGTTTAAGGCTGTTATCATTTAATGCTATGAGATACATTGT 480

QY 943 TTCTCCCTATGCCAACAACAGGTGAACAACAGTGTGTTTCTTACTGATATAAATCTTCG 1002

DB 481 TTCTCCCTATGCCAACAACAGGTGAACAACAGTGTGTTTCTTACTGATATAAATCTTCG 540

QY 1003 CTACCTGTGATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1048

DB 541 CTACCTGTGATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 586

RESULT 15

AAAL13574

ID AAL13574 standard; cDNA; 783 BP.

XX AAL13574;

XX 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 6031.

XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

XX WO200151628-A2.

XX 19-JUL-2001.

XX 10-JAN-2001; 2001WO-US000798.

XX 14-JAN-2000; 2000US-0176077P.

XX 14-MAR-2000; 2000US-0189167P.

XX 24-MAR-2000; 2000US-0192099P.

XX 29-MAR-2000; 2000US-0193480P.

XX 15-MAY-2000; 2000US-0205230P.

XX 09-JUN-2000; 2000US-0211315P.

XX 25-JUL-2000; 2000US-0220534P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer.

XX Claim 1; Page 1083; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising, treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity

XX Sequence 783 BP; 228 A; 157 C; 141 G; 248 T; 0 U; 9 Other;

XX Query Match 23.3%; Score 546.2; DB 4; Length 783;

XX Best Local Similarity 97.7%; Pred. No. 5.3e-106;

XX Matches 585; Conservative 0; Mismatches 10; Indels 4; Gaps 3;

QY 460 AAGTACTACCCCAATAAATCTTCTTATGCTCTGGAAGAGAAAGCAATGTAATCACT 519

DB 50 AGGTACTACCCCAATAAATCTTCTTATGCTCTGGAAGAGAAAGCAATGTAATCACT 109

QY 520 ATGATTTTCCACAACACAGAAATCTCCACAATTTTAACTATCTGATCTCTTCT 579

Db	110	ATGATTTATTGCAACAACCAACGAATTTCTCCACAACAATTTTAAAGTAATCTGATCCTCTTCT	169
Qy	580	TGGAGAAATCTGTACCTAAATAGTTTTTCTTATGAATGTTATTACTACTGGTATAAATC	639
Db	170	TGGAGAAATCTGTACCTAAATAGTTTTTCTTATGAATGTTATTACTACTGGTATAAATC	229
Qy	640	AAATTTCTATAAATTTCTCTACTTAAAGTCTTAAAGAACTGGGTTCTTCTCTTGATGTTATTTC	699
Db	230	AAATTTCTATAAATTTCTCTACTTAAAGTCTTAAAGAACTGGGTTCTTCTCTTGATGTTATTTC	289
Qy	700	ATGTTTCAGAAAGGAACCAACTTTTACTCTTTTATAGGACAATTCCTAGAACTCTATAGTACT	759
Db	290	ATGTTTCAGAAAGGAACCAACTTTTACTCTTTTATAGGACAATTCCTAGAACTCTATAGTACT	349
Qy	760	ATCAGAGATATATTGCTTTAAATATATTTTGGTTATTTTGAATACAGACATTTGGCTCC	819
Db	350	ATCAGAGATATATTTTGGCTTTAAATATATTTTGGTTATTTTGAATACAGACATTTGGCTCC	409
Qy	820	AAATTTTCACTCTTTGCAACAATAGTATGACTTTTCACTAGAACTTCTCAACATTTGGGAAC	879
Db	410	AAATTTTCACTCTTTGCAACAATAGTATGACTTTTCACTAGAACTTCTCAACATTTGGGAAC	469
Qy	880	TTTGCAATATGAGCATCATATGTTTAAAGCTGTATCATTTAAATGCTATGAGATACATT	939
Db	470	TTTGCAATATGAGCATCATATGTTTAAAGCTGTATCATTTAAATGCTATGAGATACATT	529
Qy	940	GTTTCTTCCCTATGCCAAACAGGTGAAC - AAACGTAGTTTGTTTTTTACTGATACATTAAAT	997
Db	530	GTTTCTTCCCTATGCCAAACAGGTGAACCAACGTAGTTGTTTTTACTGATACATTAAAT	589
Qy	998	GTTGGCTACTCTGTATTTTATAGTATGACATGTCAG - AAAAGGCAAGCAAAATGCC	1055
Db	590	GTTGGCT - CCCGNGATTTTATAGTATGACATGTCANAAAAAGGCAAGCAAAATGGC	647

Search completed: April 25, 2004, 08:46:21
Job time : 627.343 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2004, 23:54:27 ; Search time 6177.51 Seconds
(without alignments)
16460.143 Million cell updates/sec

Title: US-10-051-835-25
Perfect score: 2346
Sequence: 1 gccagaggggaaaaaagag.....aaacatttgatagttacat 2346

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.pa.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.ats.*
- 12: gb.sv.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.ats.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rtd.*
- 36: em.htg_nam.*
- 37: em.htg_vrt.*
- 38: em.sy.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	2332.4	99.4	64196	9	AC110804	AC110804 Homo sapi
C 2	2304.4	98.2	2321	6	AX281638	AX281638 Sequence
C 3	1613	68.8	1632	9	HSN806550	BS38293 Homo sapi
C 4	1599.4	68.2	1626	6	BD270271	BD270271 Homo nuc
C 5	1599.4	68.2	1626	6	AX035207	AX035207 Sequence
C 6	1599.4	68.2	1626	6	AX342061	AX342061 Sequence
C 7	1599.4	68.2	1626	6	AX342121	AX342121 Sequence
C 8	1599.4	68.2	1626	6	AX441124	AX441124 Sequence
C 9	582.8	24.8	586	6	AX351446	AX351446 Sequence
C 10	581.4	24.8	675	6	AR273526	AR273526 Sequence
C 11	581.4	24.8	675	6	AR277107	AR277107 Sequence
C 12	581.4	24.8	675	6	AX368559	AX368559 Sequence
C 13	568	24.2	586	6	AX341176	AX341176 Sequence
C 14	551.4	23.5	665	6	AR407382	AR407382 Sequence
C 15	401.8	17.1	2000	9	AB040375	AB040375 Homo sapi
C 16	323	13.8	323	6	AX333323	AX333323 Sequence
C 17	323	13.8	323	6	AX407813	AX407813 Sequence
C 18	322.2	13.7	324	6	AX341711	AX341711 Sequence
C 19	321.2	13.7	324	6	AX341282	AX341282 Sequence
C 20	321	13.7	321	6	AX322023	AX322023 Sequence
C 21	319.6	13.6	324	6	AX341552	AX341552 Sequence
C 22	273.2	11.6	277	6	AX341912	AX341912 Sequence
C 23	243.6	10.4	247	6	AX340305	AX340305 Sequence
C 24	242.6	10.3	247	6	AX341879	AX341879 Sequence
C 25	242.6	10.3	247	6	AX341931	AX341931 Sequence
C 26	242.4	10.3	244	6	AX351360	AX351360 Sequence
C 27	229.6	9.8	248	6	AX341175	AX341175 Sequence
C 28	109.8	4.7	204366	2	AC118181	AC118181 Rattus no
C 29	109.8	4.7	223893	2	AC106406	AC106406 Rattus no
C 30	93.4	4.0	196358	2	AC101992	AC101992 Mus muscu
C 31	69.8	3.0	67370	3	PFMALP3	AL031746 Plasmodiu
C 32	64	2.7	309	6	AX341780	AX341780 Sequence
C 33	63	2.7	1192	9	HSR323759	AX323759 Homo sapi
C 34	62.4	2.7	1141	6	AX083744	AX083744 Sequence
C 35	61.8	2.6	13138	3	AC116305_4	Continuation (5 of
C 36	61.8	2.6	110000	3	AC116305_3	Continuation (4 of
C 37	61.6	2.6	349751	3	PFMALP3	AL035476 Plasmodiu
C 38	61	2.6	8056	6	AX599046	AX599046 Sequence
C 39	60.4	2.6	349751	3	PFMALP3	AL035476 Plasmodiu
C 40	60.2	2.6	164900	2	BS571776	BS571776 panio rer
C 41	60.2	2.6	349880	6	AX344559	AX344559 Sequence
C 42	59.8	2.5	89665	9	AC063976	AC063976 Homo sapi
C 43	59.8	2.5	163660	2	AC046165	AC046165 Homo sapi
C 44	59.6	2.5	250029	3	AE014839	AE014839 Plasmodiu
C 45	59.2	2.5	700	6	AX183285	AX183285 Sequence

ALIGNMENTS

RESULT 1
AC110804/c
LOCUS AC110804 64196 bp DNA linear PRI 30-MAR-2002
DEFINITION Homo sapiens BAC clone RP11-725C19 from 4, complete sequence.
ACCESSION AC110804
VERSION AC110804.3 GI:19570222
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 64196)
AUTHORS Sulston,J.E. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)


```

MEDLINE
PUBMED
REFERENCE
2 (bases 1 to 64196)
Peatman,C., Kozlowicz,A. and Dignan,G.
The sequence of Homo sapiens BAC clone RP11-725C19
Unpublished (2001)
3 (bases 1 to 64196)
Waterston,R.H.
Direct Submission
Submitted (15-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 64196)
Waterston,R.H.
Direct Submission
Submitted (06-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 64196)
Waterston,R.H.
Direct Submission
Submitted (21-MAR-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6 (bases 1 to 64196)
Waterston,R.
Direct Submission
Submitted (30-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Mar 21, 2002 this sequence version replaced gi:19073866.
-----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
-----
Summary Statistics
-----
Center project name: H_NH0725C19
-----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPI1-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-8P17, 2000 bp overlap.
Actual start of this clone is at base position 185268 of
RP11-393A12; actual end is at base position 64196 of RP11-725C19.

```

```

FEATURES
    source
        Polymorphisms exist between AC013365 and AC110804.
    Location/Qualifiers
        1..64196
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="4"
            /map="4"
            /clone="RP11-725C19"
            /clone_lib="RPI1-11"
            /rpt_family="AT_rich"
            /rpt_family="AT_rich"
            /rpt_family="Alu"
            /note="match to EST BG257959 (NID:G12767775)"
        1..39
            /rpt_family="AT_rich"
        162..182
            /rpt_family="AT_rich"
        312..607
            /rpt_family="Alu"
        369..505
            /note="match to EST BG257959 (NID:G12767775)"
        877..909
            /rpt_family="AT_rich"
        1113..1536
            /rpt_family="L2"
        2400..2427
            /rpt_family="AT_rich"
        2941..2984
            /rpt_family="L2"
        3413..3435
            /rpt_family="(TA)n"
        3829..4216
            /note="match to EST H06135 (NID:G869687) Y177b12.s1"
        3829..4170
            /note="match to EST B1497113 (NID:G15336457)"
        3831..4157
            /note="match to EST AA639698 (NID:G2563477) np60f09.s1"
        3835..4360
            /note="match to EST AW237145 (NID:G6569534) xm66b03.x1"
        3836..4376
            /note="similar to EST BM791320 (NID:G19139552)"
        3837..4405
            /note="similar to Bos taurus EST B1539797 (NID:G15380907)"
        3844..4610
            /note="match to EST B1857800 (NID:G15998547)"
        3933..4394
            /note="match to EST A1694383 (NID:G4971723) wd83b04.x1"
        4105..4843
            /note="match to EST BG177619 (NID:G12684322)"
        4137..4527
            /note="match to EST A1267688 (NID:G3886855) aq93c10.x1"
        4153..4433
            /note="match to EST C00538 (NID:G1432768)"
        4233
            /note="match to EST B1497113 (NID:G15336457)"
        4359..4746
            /note="match to EST AA227760 (NID:G1849333) zr56a07.r1"
        4404..5178
            /note="match to EST BG291471 (NID:G13049386)"
        4412..4619
            /note="similar to Homo sapiens EST BF154977 (NID:G11050160)"
        4657..5028
            /note="match to EST R20307 (NID:G774941) YG35a09.r1"
        4727..5195
            /note="similar to Bos taurus EST AW669076 (NID:G7525590)"
        4740..5803
            /note="similar to EST BM547723 (NID:G18781736)"
        4859..5265
            /note="match to EST AW901105 (NID:G8065310)"
        5002..5687
            /note="match to EST BG254117 (NID:G12763933)"
        5004..5213
            /note="match to EST BF084219 (NID:G10878049)"
        5074..5447
            /note="match to EST N80420 (NID:G1243121) Y292e06.r1"

```

misc_feature 5255..6002 /note="match to EST BI869486 (NID:g16043159)" 480
misc_feature 5442..6151 /note="match to EST AW961203 (NID:g8150887)" 5701
misc_feature 5497..5830 /note="match to EST AL540897 (NID:g4458270)" 540
misc_feature 5782..6124 /note="match to EST AA309186 (NID:g1961511)" 5641
misc_feature 5875..6701 /note="match to EST BI334972 (NID:g15019629)" 600
misc_feature 5995..6331 /note="match to EST AA471165 (NID:g2199022)" 5581
misc_feature 6000..6124 /note="match to EST AW996896 (NID:g8257130)" 660
misc_feature 6032..6836 /note="similar to EST BM474943 (NID:g18523985)" 5521
misc_feature 6140..6518 /note="match to EST AW996897 (NID:g8257131)" 720
misc_feature 6444..6885 /note="match to EST AA375957 (NID:g2028275)" 5461
repeat_region 6878..7187 /rpt_family="Alu" 780
misc_feature 6899..7362 /note="match to EST AA488687 (NID:g2216118) ab38f03.sl" 5401
misc_feature 6970..7746 /note="match to EST BE739933 (NID:g10153837)" 840
misc_feature 7099..7815 /note="match to EST BF381953 (NID:g11363256)" 5341
misc_feature 7124..7125 /note="match to EST AW996897 (NID:g8257131)" 900
misc_feature 7174..7495 /note="match to EST BF927528 (NID:g12325656)" 5281
misc_feature 7253..7884 /note="similar to Homo sapiens EST BG499179 (NID:g13460696)" 960
misc_feature 7568..7949 /note="match to EST AA307001 (NID:g1959330)" 5221
misc_feature 7613..8332 /note="similar to EST BM706727 (NID:g19019985)" 1020
repeat_region 7836..8011 /rpt_family="AchHob" 5161

Query Match 99.4%; Score 2332.4; DB 9; Length 64196;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2344; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GCCAGAGGGGAAAAAGAGTAATGCACAGGTATCTCTTTTCAGTGGTGAAGTATTTT 60
DB 6179 GCCAGAGGGGAAAAAGAGTAATGCACAGGTATCTCTTTTCAGTGGTGAAGTATTTT 6120

QY 61 GAGTACTTGTGTGACAGGTAATATACAGCATCTCTGCGGAACCTATTAGGCCCTT 120
DB 6119 GAGTACTTGTG- TGTGACAGGTAATATACAGCATCTCTTTCAGTGGTGAAGTATTTT 6061

QY 121 TGCATGTTTAAAGCTGTATATTTTGGTGTGTGAGTGGTCTGACTTAAATGCTATTAT 180
DB 6060 TGCATGTTTAAAGCTGTATATTTTGGTGTGTGAGTGGTCTGACTTAAATGCTATTAT 6001

QY 181 AAAATTTAGACATCAAAATTTTCTCTACTACTAATCTTATAGATGATATCTGGAAGCAC 240
DB 6000 AAAATTTAGACATCAAAATTTTCTCTACTACTAATCTTATAGATGATATCTGGAAGCAC 5941

QY 241 AGTCATATACACTGGGAGCAATGCAATGTGTGTACCTGGTCTGAGTTTGAACCTGCT 300
DB 5940 AGTCATATACACTGGGAGCAATGCAATGTGTGTACCTGGTCTGAGTTTGAACCTGCT 5881

QY 301 TATTTCAAAAGATTTCTGAAATTAATTTTCCCTAGAAATTTCTCTTCATCTCCAAAGTACA 360
DB 5880 TATTTCAAAAGATTTCTGAAATTAATTTTCCCTAGAAATTTCTCTTCATCTCCAAAGTACA 5821

QY 361 AACATCTTTGAAGATGAACAGATTTGTTCCCATGAATGTATGTCTCATCTGACTAGA 420
DB 5820 AACATCTTTGAAGATGAACAGATTTGTTCCCATGAATGTATGTCTCATCTGACTAGA 5761

QY 421 AACGATCTATGTAAATGACTGTATATGAATTTTCAAGTACTACCCCAATAACTT 480
DB 5760 AACGATCTATGTAAATGACTGTATATGAATTTTCAAGTACTACCCCAATAACTT 5701

QY 481 TCTTATTGCTCTGAAAGAGAAAAGCAATGTAATCACTATGATTTGCACAAAACAAC 540
DB 5700 TCTTATTGCTCTGAAAGAGAAAAGCAATGTAATCACTATGATTTGCACAAAACAAC 5641

QY 541 AGAATTTCTCCAACTTTTAAGTAACTGATCTCTCTTTGGAGAAAATTTGTTACCTAAT 600
DB 5640 AGAATTTCTCCAACTTTTAAGTAACTGATCTCTCTTTGGAGAAAATTTGTTACCTAAT 5581

QY 601 AGTTTTTCTTTAGGAATGTTTATTACTACTAGGTATAAATCAAAATTTCTATAAAATTTCTCTAC 660
DB 5580 AGTTTTTCTTTAGGAATGTTTATTACTACTAGGTATAAATCAAAATTTCTATAAAATTTCTCTAC 5521

QY 661 TTAAGTCTTTAAGAACTGGGTCTCTCTTTGATGTTTATTCAATGTTCAAGAGAAAACAACA 720
DB 5520 TTAAGTCTTTAAGAACTGGGTCTCTCTTTGATGTTTATTCAATGTTCAAGAGAAAACAACA 5461

QY 721 CTTTACTCTTTTAAAGCAATTTCTAGAACTATAGTAGTATCAGGATATATTTTGCTTTA 780
DB 5460 CTTTACTCTTTTAAAGCAATTTCTAGAACTATAGTAGTATCAGGATATATTTTGCTTTA 5401

QY 781 AAATATATTTTGGTATTTTGAATACAGCAATTTGGCTCCAAATTTTCACTTTTGCACAAT 840
DB 5400 AAATATATTTTGGTATTTTGAATACAGCAATTTGGCTCCAAATTTTCACTTTTGCACAAT 5341

QY 841 AGTATGACTTTTCACTAGAACTTTCTCAACATTTTGGGAATTTTCAAAATATAGCAATCATYA 900
DB 5340 AGTATGACTTTTCACTAGAACTTTCTCAACATTTTGGGAATTTTCAAAATATAGCAATCATYA 5281

QY 901 TGTGTTAAGGCTGTATCAATTTAATGCTATGAGATACATTTGTTTTCTCTCCCTATGCCAACA 960
DB 5280 TGTGTTAAGGCTGTATCAATTTAATGCTATGAGATACATTTGTTTTCTCTCCCTATGCCAACA 5221

QY 961 GGTGAACAAACGCTAGTTGTTTTTACTGATCTAAATTTTGGCTACTGCTGATTTTATAG 1020
DB 5220 GGTGAACAAACGCTAGTTGTTTTTACTGATCTAAATTTTGGCTACTGCTGATTTTATAG 5161

QY 1021 TATGCATCTCAGAAAAGGCAAGCAAAATGGCTCTGTACTGTAATCTTCGGCAAC 1080
DB 5160 TATGCATCTCAGAAAAGGCAAGCAAAATGGCTCTGTACTGTAATCTTCGGCAAC 5101

QY 1081 TTATTGGGTCTTCAATTTTCTGACAGCAGGATTTGCTCAATTTTGTAGAGCTTTGGTA 1140
DB 5100 TTATTGGGTCTTCAATTTTCTGACAGCAGGATTTGCTCAATTTTGTAGAGCTTTGGTA 5041

QY 1141 GAATGGATTTACATGTTAGTATGACCTGGTAGAAATGGTTTTTGTATTGACTCAGAAT 1200
DB 5040 GAATGGATTTACATGTTAGTATGACCTGGTAGAAATGGTTTTTGTATTGACTCAGAAT 4981

QY 1201 TCATCTCAGGATGAATCTTTTATGTCTTTTATGTATGTATGTATGTATGTATGTATGTATGTAT 1260
DB 4980 TCATCTCAGGATGAATCTTTTATGTCTTTTATGTATGTATGTATGTATGTATGTATGTATGTAT 4921

QY 1261 AAGATGGTTTTAGAAAGCTTTGTCTAAAAATTTGGCCTAGGAATGGTAATCTCAATTTTCA 1320
DB 4920 AAGATGGTTTTAGAAAGCTTTGTCTAAAAATTTGGCCTAGGAATGGTAATCTCAATTTTCA 4861

QY 1321 GTTCCCAAGGGGTAGAAAAATAATATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
DB 4860 GTTCCCAAGGGGTAGAAAAATAATATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4801

QY 1381 GTACTTCTATGAATGTATTTTAAATATTTTTCATATTTCTGTGCAAGCATTTTATAATTTT 1440
DB 4800 GTACTTCTATGAATGTATTTTAAATATTTTTCATATTTCTGTGCAAGCATTTTATAATTTT 4741

QY 1441 CAACAAGTGGAGTCCATTTAGCCAGTGGGAAGCTTTGGAACTTCAGGTTACCCCTGAAG 1500
DB 4740 CAACAAGTGGAGTCCATTTAGCCAGTGGGAAGCTTTGGAACTTCAGGTTACCCCTGAAG 4681

1501 GATATCTGGCAGCCATCTCTTTGATCTGTCGCTTAACTGTAAATTTATAGACCACTAAA 1560
|
|
|
Db GATATCTGGCAGCCATCTCTTTGATCTGTCGCTTAACTGTAAATTTATAGACCACTAAA 4621
|
|
|
Qy TCCTAACTTGGATCTGGATGCAATAGTATGACCTCTGACATTTCCAGAAATTTCCAG 1620
|
|
|
Db TCCTAACTTGGATCTGGATGCAATAGTATGACCTCTGACATTTCCAGAAATTTCCAG 4561
|
|
|
Qy GGATCGTGGTCTGCTAGTATGAAACACAAAGACAGAGATCCAGCTGAAA 1680
|
|
|
Db GGATCGTGGTCTGCTAGTATGAAACACAAAGACAGAGATCCAGCTGAAA 4501
|
|
|
Qy GAGTATCTCAATATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCT 1740
|
|
|
Db GAGTATCTCAATATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCT 4441
|
|
|
Qy GCACTGTGATCATGAACCTTAGTAGAGGATGCTGTATTTATTAACAAATTTAATACA 1800
|
|
|
Db GCACTGTGATCATGAACCTTAGTAGAGGATGCTGTATTTAATTAACAAATTTAATACA 4381
|
|
|
Qy ATGCTTTACATGATTAATCTTAAGAGCAAACTGCAATTTATTTCTGATCCACAT 1860
|
|
|
Db ATGCTTTACATGATTAATCTTAAGAGCAAACTGCAATTTATTTCTGATCCACAT 4321
|
|
|
Qy TCCAATCATATTAGAACTAAGATATTTATCTATGAAGATATAAATGGTCCAGAGACTT 1920
|
|
|
Db TCCAATCATATTAGAACTAAGATATTTATCTATGAAGATATAAATGGTCCAGAGACTT 4261
|
|
|
Qy TCATCTGTGATTCGGTCTGCTAGGCTTCTAGGCTTCTAGGCTTCTAGGCTTCTAGGCT 1980
|
|
|
Db TCATCTGTGATTCGGTCTGCTAGGCTTCTAGGCTTCTAGGCTTCTAGGCTTCTAGGCT 4201
|
|
|
Qy ATATGTGAATATAAATGGATCTCTATAGCTAAATGAGTTCCTCTGGGAGAGTCTG 2040
|
|
|
Db ATATGTGAATATAAATGGATCTCTATAGCTAAATGAGTTCCTCTGGGAGAGTCTG 4141
|
|
|
Qy GTACTGCAATCACATGCCAGATGGTGTATGCGCTATTTGTAAGTAAAGTGAAGA 2100
|
|
|
Db GTACTGCAATCACATGCCAGATGGTGTATGCGCTATTTGTAAGTAAAGTGAAGA 4081
|
|
|
Qy TGCTATGAAGTAAAGTGGTGTGTTTTCATCTATGGAACCTCTGATGATGCTGCTTTG 2160
|
|
|
Db TGCTATGAAGTAAAGTGGTGTGTTTTCATCTATGGAACCTCTGATGATGCTGCTTTG 4021
|
|
|
Qy TATGGAATAAATTTGGTGCATATGATGTCAATTCGATTTGAATTTGAATTTGG 2220
|
|
|
Db TATGGAATAAATTTGGTGCATATGATGTCAATTCGATTTGAATTTGAATTTGG 3961
|
|
|
Qy TTGTAATTTATATGATATACCTGTACGCTTCTAGTGTCTTCAACCATTTTATACCAT 2280
|
|
|
Db TTGTAATTTATATGATATACCTGTACGCTTCTAGTGTCTTCAACCATTTTATACCAT 3901
|
|
|
Qy TTTTGTACATTTTACTTGAATAATTTTAAATGGAATTTTAAATGGAATTTTAAATGGA 2340
|
|
|
Db TTTTGTACATTTTACTTGAATAATTTTAAATGGAATTTTAAATGGAATTTTAAATGGA 3841
|
|
|
Qy TTACAT 2346
|
|
|
Db TTACAT 3835

RESULT 2
AX281638
LOCUS AX281638 2321 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 47 from Patent WO0177389.
ACCESSION AX281638
VERSION AX281638.1 GI:16608889
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE

Shiffman,D., Sonoyvi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,
Mikita,T. and Tai,J.
Genes expressed in foam cell differentiation
Patent: WO 0177389-A 47 18-OCT-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1..2321
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 200578.1"

ORIGIN
Query Match 98.2%; Score 2304.4; DB 6; Length 2321;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2316; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Qy 28 CAGTATCTCTTTGAGTGGTGTGATTTAGTACCTTGTGAGTGGTGTGATTTAT 87
|
|
|
Db 4 CAGTATCTCTTTGAGTGGTGTGATTTAGTACCTTGTGAGTGGTGTGATTTAT 62
|
|
|
Qy 88 TACAGCATCTTGGGAAAACCTATTAGGCTTTGATGTTTAAAGCTGTATATTTGTTG 147
|
|
|
Db 63 TACAGCATCTTGGGAAAACCTATTAGGCTTTGATGTTTAAAGCTGTATATTTGTTG 122
|
|
|
Qy 148 GGTGTGAGTGGTGTGATTTAAATGTTATTAATAATTTAGACATCAAAATTTCTTACT 207
|
|
|
Db 123 GGTGTGAGTGGTGTGATTTAAATGTTATTAATAATTTAGACATCAAAATTTCTTACT 182
|
|
|
Qy 208 AACTAACTTTATAGTGCATCTTGAAGCAGATCATATCACCTGGGAGCAATGCA 267
|
|
|
Db 183 AACTAACTTTATAGTGCATCTTGAAGCAGATCATATCACCTGGGAGCAATGCA 242
|
|
|
Qy 268 ATGGGTTCACCTGCTCCTAGTGTGAACTGTCTTTATTTCAAAAGATTTGAAATTT 327
|
|
|
Db 243 ATGGGTTCACCTGCTCCTAGTGTGAACTGTCTTTATTTCAAAAGATTTGAAATTT 302
|
|
|
Qy 328 TTCCCTAGAAATTTCTCTTCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 387
|
|
|
Db 303 TTCCCTAGAAATTTCTCTTCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362
|
|
|
Qy 388 GTTCCCATGAATGTATGCTCATCTCGACTAGAAACGATCTATGTTAAATGACTGTAT 447
|
|
|
Db 363 GTTCCCATGAATGTATGCTCATCTCGACTAGAAACGATCTATGTTAAATGACTGTAT 422
|
|
|
Qy 448 ATGAATTTTCAAGTACTACCCCAATTAATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 507
|
|
|
Db 423 ATGAATTTTCAAGTACTACCCCAATTAATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 482
|
|
|
Qy 508 ATGTAATCACTATGATTTATGCAAAACCAAGAAATTTCTCCAAACAAATTTTAAATTAATC 567
|
|
|
Db 483 ATGTAATCACTATGATTTATGCAAAACCAAGAAATTTCTCCAAACAAATTTTAAATTAATC 542
|
|
|
Qy 568 TGATCTCTCTTGGAGAAATTTGTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 627
|
|
|
Db 543 TGATCTCTCTTGGAGAAATTTGTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602
|
|
|
Qy 628 CTGGTATAAATCAAAATTTCTATAAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 687
|
|
|
Db 603 CTGGTATAAATCAAAATTTCTATAAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 662
|
|
|
Qy 688 TTGATGTTATTCATGTTGAGAAAGAAACCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 747
|
|
|
Db 663 TTGATGTTATTCATGTTGAGAAAGAAACCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 722
|
|
|
Qy 748 ATCTATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 807
|
|
|
Db 723 ATCTATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 782
|
|
|
Qy 808 GACATGCTCAAAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 867
|
|
|
Db 783 GACATGCTCAAAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 842

```
QY 868 ACATTTGGGAACCTTTGCAAAATATGAGCATCATATGTGTAAAGGCTGTATCATTTAAAGCT 927
Db 843 ACATTTGGGAACCTTTGCAAAATATGAGCATCATATGTGTAAAGGCTGTATCATTTAAAGCT 902
QY 928 ATGAGATACATTTGTTTCTCCCTATGCGCAAAACAGAGTGAACAAACGCTAGTGTGTTTAACT 987
Db 903 ATGAGATACATTTGTTTCTCCCTATGCGCAAAACAGAGTGAACAAACGCTAGTGTGTTTAACT 962
QY 988 GATACATAATGTGGCTACCTGTGTATTTTATAGTATGCAATGTGCAGAAAAGGCAAGAC 1047
Db 963 GATACATAATGTGGCTACCTGTGTATTTTATAGTATGCAATGTGCAGAAAAGGCAAGAC 1022
QY 1048 AAATGGCCCTCTGTACTCAATACCTTGGGCAAACTTATTTGGGCTCTTCAATTTCTGACGAC 1107
Db 1023 AAATGGCCCTCTGTACTCAATACCTTGGGCAAACTTATTTGGGCTCTTCAATTTCTGACGAC 1082
QY 1108 AGGATTTGACTCAATATTTGTAGAGCTTGGTAGAATGATATGATGATGATGATGATGATGAT 1167
Db 1083 AGGATTTGACTCAATATTTGTAGAGCTTGGTAGAATGATATGATGATGATGATGATGATGAT 1142
QY 1168 GGTAGAAATGGTTTATTTAGTTATTTGACTCAGAACTCATCTCAGGATGAATCTTTTATGTCT 1227
Db 1143 GGTAGAAATGGTTTATTTAGTTATTTGACTCAGAACTCATCTCAGGATGAATCTTTTATGTCT 1202
QY 1228 TTTTATGTGAGCATATCTGAATTTACTTTTAAAGATGGTTTGAAGAGCTTTGTCTAA 1287
Db 1203 TTTTATGTGAGCATATCTGAATTTACTTTTAAAGATGGTTTGAAGAGCTTTGTCTAA 1262
QY 1298 AAATTTGGCCCTAGGAATGTAACTTCAATTTTCAAGTTGCAAGGGGTAGAAAATAATATG 1347
Db 1263 AAATTTGGCCCTAGGAATGTAACTTCAATTTTCAAGTTGCAAGGGGTAGAAAATAATATG 1322
QY 1348 TGTGTTGTTATGTTTATGTTAAATATTTATTTAGGTTACTATCTATGAATGTTATTTAAATAT 1407
Db 1323 TGTGTTGTTATGTTTATGTTAAATATTTATTTAGGTTACTATCTATGAATGTTATTTAAATAT 1382
QY 1408 TTTTCATATCTGTGACAGCAATTTATTTTGAACAGTGGAGTCCATTTAGCCCACT 1467
Db 1383 TTTTCATATCTGTGACAGCAATTTATTTTGAACAGTGGAGTCCATTTAGCCCACT 1442
QY 1468 GGGAAAGCTTCTGGAACCTCAGGTTTACCTTTGAAGGATATGCTGGCAGCAATCTCTTTGATC 1527
Db 1443 GGGAAAGCTTCTGGAACCTCAGGTTTACCTTTGAAGGATATGCTGGCAGCAATCTCTTTGATC 1502
QY 1528 TGTGCTTAACTGTTATTTATGACAGTAAATCCCTAACTTGAATGGAATGCAATTA 1587
Db 1503 TGTGCTTAACTGTTATTTATGACAGTAAATCCCTAACTTGAATGGAATGCAATTA 1562
QY 1588 GTTATGACCTTGTACCATTTCCAGAAATTTTCAAGGATGATGCTGGGTTTGGTCTAGTATG 1647
Db 1563 GTTATGACCTTGTACCATTTCCAGAAATTTTCAAGGATGATGCTGGGTTTGGTCTAGTATG 1622
QY 1648 AAAACAAAGAACAGAGATCCAGTGAAGAGTGAATCTCAATATCTTAACTAACT 1707
Db 1623 AAAACAAAGAACAGAGATCCAGTGAAGAGTGAATCTCAATATCTTAACTAACT 1682
QY 1708 GGTCTCTCAACTCAAGCAGAGTTTCTTCACTCTGGCACTGTGATCATGAACTTAGTAGAG 1767
Db 1683 GGTCTCTCAACTCAAGCAGAGTTTCTTCACTCTGGCACTGTGATCATGAACTTAGTAGAG 1742
QY 1768 GGGATTTGTGTATTTTATACAAATTAATAAATGTTCTTAATGATGATAAATTTCTTAA 1827
Db 1743 GGGATTTGTGTATTTTATACAAATTTAATAAATGTTCTTCAATGATAAATTTCTTAA 1802
QY 1828 GAGCAAACTGCAATTTTATTTCTGATCCACATTTCCAACTATTTAGAACTAAAGATATTT 1887
Db 1803 GAGCAAACTGCAATTTTATTTCTGATCCACATTTCCAACTATTTAGAACTAAAGATATTT 1862
QY 1888 ATCTATGAGATATAAATGGTGACAGAGACTTTTCACTGTGGATTTGCGTTTCTTTAG 1947
Db 1863 ATCTATGAGATATAAATGGTGACAGAGACTTTTCACTGTGGATTTGCGTTTCTTTAG 1922
QY 1948 GGTCTCTAGCACTGATCCCTGCAAGCATGTGATATGTGAATAAATGGATTTCTTCTA 2007
```

```
Db 1923 GGTCTCTAGCACTGATGCTCTGACCAAGCATGTGATATGTGAATAAATGGATTTCTTCTA 1982
QY 2008 TAGCTAAATGAGTTCCCTCTGCGGAGAGTTCTGGTACTGCAATCAATGCCAGATGGTG 2067
Db 1983 TAGCTAAATGAGTTCCCTCTGCGGAGAGTTCTGGTACTGCAATCAATGCCAGATGGTG 2042
QY 2068 TTTATGGGCTATTGTTGTAAGTAAGTGGTAAGATGCTATGAAGTAAGTGTGTTGTTTTC 2127
Db 2043 TTTATGGGCTATTGTTGTAAGTAAGTGGTAAGATGCTATGAAGTAAGTGTGTTGTTTTC 2102
QY 2128 ATCTTAGGAACTCTTTGATGCAATGCTTTTGTATGGAATAAATTTGGTGCATATGA 2187
Db 2103 ATCTTAGGAACTCTTTGATGCAATGCTTTTGTATGGAATAAATTTGGTGCATATGA 2162
QY 2188 TGTCAATCAACTTTGCAATTTGAATTTGGTGTATTTATATATATATATACCTGTCA 2247
Db 2163 TGTCAATCAACTTTGCAATTTGAATTTGGTGTATTTATATATATATATACCTGTCA 2222
QY 2248 CGTCTTAGTGTCTCAACCACTTTATTAACCACTTTTGTATATATATATATACCTGTCA 2307
Db 2223 CGTCTTAGTGTCTCAACCACTTTTATAACCACTTTTGTATATATATATATACCTGTCA 2282
QY 2308 TTTAATGGAATTTTAAATAAACAATTTGATGTTTACAT 2346
Db 2283 TTTAATGGAATTTTAAATAAACAATTTGATGTTTACAT 2321

RESULT 3
HSM806550 1632 bp mRNA linear PRI 18-JUN-2003
LOCUS Homo sapiens mRNA; cDNA DKFZp686F09157 (from clone DKFZp686F09157).
DEFINITION BX538293
ACCESSION BX538293.1 GI:31874775
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 1632)
AUTHORS Lauber, J., Bahr, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobo, G., Han, M. and Wiemann, S.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; being
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp686F09157) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.

FEATURES
Location/Qualifiers
1..1632
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="4Q26.3"
/clone="DKFZp686F09157"
/tissue_type="human colon endothel primary cell culture"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
polya_signal 1591..1596
polya_site 1609
ORIGIN
Query Match 68.8%; Score 1613; DB 9; Length 1632;
Best Local Similarity 100.0%; Pred. No. 5.9e-279;
Matches 1613; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	734	GGACAAATTCCTAGAAATCTATAGTAGTATCAGGATATATTTTGCCTTTAAATATATTTTGG	793
Db	1	GGACAAATTCCTAGAAATCTATAGTAGTATCAGGATATATTTTGCCTTTAAATATATTTTGG	60
Qy	794	TTATTTTGAATACAGACATTTGGCTCCAAATTTTCACTTTTGCACAATAGTATGACTTTTC	853
Db	61	TTATTTTGAATACAGACATTTGGCTCCAAATTTTCACTTTTGCACAATAGTATGACTTTTC	120
Qy	854	ACTAGAACTTCTCAACATTTGGGAACCTTTGCAAAATNTGAGCATCATATGTGTTAAGCTG	913
Db	121	ACTAGAACTTCTCAACATTTGGGAACCTTTGCAAAATNTGAGCATCATATGTGTTAAGCTG	180
Qy	914	TATCAATTTAATGCTATGAGATACATTTGTTTCTCCCTATGCAAAACAGGTGAACAAACGT	973
Db	181	TATCAATTTAATGCTATGAGATACATTTGTTTCTCCCTATGCAAAACAGGTGAACAAACGT	240
Qy	974	AGTTGTTTTTTTACTGATACTAAATGTTGGCTACCTGTGATTTTATAGTATGCACATGTCA	1033
Db	241	AGTTGTTTTTTTACTGATACTAAATGTTGGCTACCTGTGATTTTATAGTATGCACATGTCA	300
Qy	1034	GA AAAAGGCAAGA CA AATGGGCTCTTGCTACTGAAATCTTGGCAAACTTATTTGGGCTTC	1093
Db	301	GA AAAAGGCAAGA CA AATGGGCTCTTGCTACTGAAATCTTGGCAAACTTATTTGGGCTTC	360
Qy	1094	ATTTTCTGACAGACAGATTTGACTCAATATTTCTGAGAGCTTGGTAGAATGATACAT	1153
Db	361	ATTTTCTGACAGACAGATTTGACTCAATATTTCTGAGAGCTTGGTAGAATGATACAT	420
Qy	1154	GGTAGTGATGCCTGGTAGAAATGGTTTTTATGTTATTGACTCAGAAATCATCTCAGGATG	1213
Db	421	GGTAGTGATGCCTGGTAGAAATGGTTTTTATGTTATTGACTCAGAAATCATCTCAGGATG	480
Qy	1214	AATCTTTTTATGCTTTTTTATTGTAAGCATATCTGAAATTTACTTTTATAAAGATGGTTTTAG	1273
Db	481	AATCTTTTTATGCTTTTTTATTGTAAGCATATCTGAAATTTACTTTTATAAAGATGGTTTTAG	540
Qy	1274	AAAGCTTTGCTCAAAAATTTGGGCTAGAAATGGTAACTTCAATTTTTCAGTTGCCAAGGGGT	1333
Db	541	AAAGCTTTGCTCAAAAATTTGGGCTAGAAATGGTAACTTCAATTTTTCAGTTGCCAAGGGGT	600
Qy	1334	AGAAAATAATATGCTGTGTTGTTATGTTTATGTTTAAACATATTTAGGTACTATCTATGA	1393
Db	601	AGAAAATAATATGCTGTGTTGTTATGTTTATGTTTAAACATATTTAGGTACTATCTATGA	660
Qy	1394	ATGTATTTAAATATTTTTTCATATCTGTGACAAGCATTTTAAATTTGCAACAAGTGGAGT	1453
Db	661	ATGTATTTAAATATTTTTTCATATCTGTGACAAGCATTTTAAATTTGCAACAAGTGGAGT	720
Qy	1454	CCATTTAGCCAGTGGGAAAGTCTTGGAACTCAGGTTACCTTGAAGGATATGCTGGCAG	1513
Db	721	CCATTTAGCCAGTGGGAAAGTCTTGGAACTCAGGTTACCTTGAAGGATATGCTGGCAG	780
Qy	1514	CCATCTCTTTGATCTGTCTTAACTGTATTTATAGACAGCTTAAATCCCTAACTTGGGA	1573
Db	781	CCATCTCTTTGATCTGTCTTAACTGTATTTATAGACAGCTTAAATCCCTAACTTGGGA	840
Qy	1574	TCTGGAATGCAATTAGTTATGACCTTGTAACATTTCCAGAAATTTCCAGGGGCATCGTGGGTT	1633
Db	841	TCTGGAATGCAATTAGTTATGACCTTGTAACATTTCCAGAAATTTCCAGGGGCATCGTGGGTT	900
Qy	1634	TGGTCTAGTGTGTTGAAAACAAGAAACAGAGAGATCCAGCTGAAAAGAGTGCCTCAA	1693
Db	901	TGGTCTAGTGTGTTGAAAACAAGAAACAGAGAGATCCAGCTGAAAAGAGTGCCTCAA	960
Qy	1694	TATCCTTAACCTAACTGGTCTCTCAACTCAAGCAGAGTTTCTTCACTCTGGCACTGTGATCAT	1753
Db	961	TATCCTTAACCTAACTGGTCTCTCAACTCAAGCAGAGTTTCTTCACTCTGGCACTGTGATCAT	1020
Qy	1754	GA AACTTTAGAGGGGATTTGTGTATTTTATACAAATTTTATACAAATGTCTTACATTTG	1813
Db	1021	GA AACTTTAGAGGGGATTTGTGTATTTTATACAAATTTTATACAAATGTCTTACATTTG	1080
Qy	1814	ATAAAATTTCTTAAAGAGCAAACTGCATTTTTTATTTCTGCAATCCACATTCCAATCATATTA	1873

Db	1081	ATATAATTTCTTAAGAGCAAAATCGCATTTTATTTCTGCAATCGACATTCACATCATATTA	1140
Qy	1874	GAACCTAAGATATTTATCTATGAAGATATAAATGGTGACAGAGACTTTTCATCTGTGGATT	1933
Db	1141	GAACCTAAGATATTTATCTATGAAGATATAAATGGTGACAGAGACTTTTCATCTGTGGATT	1200
Qy	1934	CGCTGTGTTTCTTAGGGTTCTTAGCACATGATCCCTGCACACAGCATGTGATATGGAATAA	1993
Db	1201	CGCTGTGTTTCTTAGGGTTCTTAGCACATGATCCCTGCACACAGCATGTGATATGGAATAA	1260
Qy	1994	AATGGATTCTTCTATAGCTAAATGAGTTCCCTCTGGGAGAGTTCTGGTACTGCAATCAC	2053
Db	1261	AATGGATTCTTCTATAGCTAAATGAGTTCCCTCTGGGAGAGTTCTGGTACTGCAATCAC	1320
Qy	2054	AATGCCAGATGGTGTATATGGGCTATTGTGTAAAGTAAAGTAAAGTAAAGTAAAGTAA	2113
Db	1321	AATGCCAGATGGTGTATATGGGCTATTGTGTAAAGTAAAGTAAAGTAAAGTAAAGTAA	1380
Qy	2114	GTGTGTTTGTTCCTCACTTATGAAGAACTCTTGATGATGTCGCTTTTGTATGGAATAAAT	2173
Db	1381	GTGTGTTTGTTCCTCACTTATGAAGAACTCTTGATGATGTCGCTTTTGTATGGAATAAAT	1440
Qy	2174	TTGTGCAATATGATGTCATTCACCTTTGATGAAATGAAATTTGGTTCCTATTATATG	2233
Db	1441	TTGTGCAATATGATGTCATTCACCTTTGATGAAATGAAATTTGGTTCCTATTATATG	1500
Qy	2234	TATTATACCTGTCACGCTTCTAGTGTCTTCAACCATTTTATAACCATTTTGTACATATT	2293
Db	1501	TATTATACCTGTCACGCTTCTAGTGTCTTCAACCATTTTATAACCATTTTGTACATATT	1560
Qy	2294	TTACTTGAAATATTTTAAATGAAATTTTAAATGAAATTTTAAATGAAATTTTAAATG	2346
Db	1561	TTACTTGAAATATTTTAAATGAAATTTTAAATGAAATTTTAAATGAAATTTTAAATG	1613
RESULT 4			
BD270271		1626 bp DNA linear	PAT 17-JUL-2003
LOCUS			
DEFINITION		Human nucleic acid sequence and protein sequence derived from	
		endothelial cell.	
ACCESSION		BD270271	
VERSION		BD270271.1	GI:33080039
KEYWORDS		JP 2002537832-A/24.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE		1. (bases 1 to 1626)	
		Thierauch,K.H., Glienke,J., Hinzmann,B. and Pilarsky,C.	
		Human nucleic acid sequence and protein sequence derived from	
		endothelial cell	
JOURNAL		Patent: JP 2002537832-A 24 12-NOV-2002;	
COMMENT		SCHERING AKTIENGESSELLSCHAFT	
		OS Homo sapiens (human)	
		PN JP 2002537832-A/24	
		PD 12-NOV-2002	
		PF 08-MAR-2000 JP 2000603357	
		PR 09-MAR-1999 DE 199 11 684 9, 01-OCT-1999 DE 199 48 679, 4 PI	
		KARL HEINZ THIERAUCH, JENS GLENKE, BERND HINZMANN, CHRISTIAN PI	
		PIIARSKY	
		PC C12N15/09, A61K31/711, A61K38/00, A61K48/00, A61P1/16, A61P3/10, PC	
		A61P9/10,	
		PC A61P9/10, A61P9/14, A61P13/12, A61P17/06, A61P19/02, A61P25/00, PC	
		A61P27/02,	
		PC A61P27/06, A61P29/00, A61P35/00, A61P41/00, A61P43/00, A61P43/00,	
		PC C07K14/47,	
		PC C07K16/18, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12P21/08, PC	
		C12N15/00,	
		PC C12N5/00, A61K37/02	
		CC Human nucleic acid sequence and protein sequence derived from	
		endothelial	
		CC cell	

ORIGIN

Query Match 68.2%; Score 1599.4; DB 6; Length 1626;
Best Local Similarity 99.9%; Pred. No. 1.6e-276;
Matches 1611; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 734 GGCAATTCCTAGAACTCTAGTAGTATCAGGATATATTTTGGTAAATATATTTGG 793
DB 1 GGCAATTCCTAGAACTCTAGTAGTATCAGGATATATTTTGGTAAATATATTTGG 60

QY 794 TTATTTTGAATACAGACATTTGGCTCCAAATTTTCATCTTTTGCAATATGATGCTTTTC 853
DB 61 TTATTTTGAATACAGACATTTGGCTCCAAATTTTCATCTTTTGCAATATGATGCTTTTC 120

QY 854 ACTAGAACTTCTCAACATTTGGGAATTTTGCATATATGACATCATCTGTTTAAAGGCTG 913
DB 121 ACTAGAACTTCTCAACATTTGGGAATTTTGCATATATGACATCATCTGTTTAAAGGCTG 180

QY 914 TATCATTTAAATGCTATGAGATACATTTGTTTCTCCCTATGCCAAACAGGTGAACAAAGT 973
DB 181 TATCATTTAAATGCTATGAGATACATTTGTTTCTCCCTATGCCAAACAGGTGAACAAAGT 240

QY 974 AGTTGTTTTTACTGATCTAAATTTGGCTACCTGATTTTATAGTATGACATGCTCA 1033
DB 241 AGTTGTTTTTACTGATCTAAATTTGGCTACCTGATTTTATAGTATGACATGCTCA 300

QY 1034 GAAAAAGGCAAGCAAAATGGCTCTTGTACTGAATACTTCGCAAACTTTATGGGCTTC 1093
DB 301 GAAAAAGGCAAGCAAAATGGCTCTTGTACTGAATACTTCGCAAACTTTATGGGCTTC 360

QY 1094 ATTTTCTGACAGACAGATTTGACTCAATTTTGTAGAGCTGCGPAGATGGATACAT 1153
DB 361 ATTTTCTGACAGACAGATTTGACTCAATTTTGTAGAGCTGCGPAGATGGATACAT 420

QY 1154 GGTAGTGACCTGGTGAATATGGTTTATTTAGTATGACTCAGAACTTCCTCAGGATG 1213
DB 421 GGTAGTGACCTGGTGAATATGGTTTATTTAGTATGACTCAGAACTTCATCTCAGGATG 480

QY 1214 AATCTTTTATGCTTTTATTTGTAAGCATCTGAATTTTACTTTTAAAGATGGTTTATG 1273
DB 481 AATCTTTTATGCTTTTATTTGTAAGCATCTGAATTTTACTTTTAAAGATGGTTTATG 540

QY 1274 AAAGCTTTGTCTAAAAATTTGGCTAGGAATGTTAACTTCATTTTCAGTTGCCAAGGGT 1333
DB 541 AAAGCTTTGTCTAAAAATTTGGCTAGGAATGTTAACTTCATTTTCAGTTGCCAAGGGT 600

QY 1334 AGAAAAATATATGTTGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATG 1393
DB 601 AGAAAAATATATGTTGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATG 660

QY 1394 ATGTATTTAAATATTTTTCATATCTGTGCAAGCATTTTATATTTTGCACAAAGTGGAGT 1453
DB 661 ATGTATTTAAATATTTTTCATATCTGTGCAAGCATTTTATATTTTGCACAAAGTGGAGT 720

QY 1454 CCATTTAGCCAGTGGGAAAGTCTTGGAATCTAGGTTACCCCTTGAAGGATATGCTGGCAG 1513
DB 721 CCATTTAGCCAGTGGGAAAGTCTTGGAATCTAGGTTACCCCTTGAAGGATATGCTGGCAG 780

QY 1514 CCATCTCTTTGATCTGCTTAACTGTAATTTATGACAGCTAAATCCCTAACTTTGA 1573
DB 781 CCATCTCTTTGATCTGCTTAACTGTAATTTATGACAGCTAAATCCCTAACTTTGA 840

QY 1574 TCTGGAATGATTTATGATCTTTGATCCATTTCCAGAAATTTCCAGGGATCTGCTGGTT 1633
DB 841 TCTGGAATGATTTATGATCTTTGATCCATTTCCAGAAATTTCCAGGGATCTGCTGGTT 899

QY 1634 TGGTCTAGTATGTAACACAGACAGATCCAGCTGCAAAAGAGTATCTCTCA 1693
DB 900 TGGTCTAGTATGTAACACAGACAGATCCAGCTGCAAAAGAGTATCTCTCA 959

QY 1694 TATCCTAACTAACTGGTCTTCACTCAAGCAGAGTTTCTTCACTCTGGCACTGTGATCAT 1753
DB 960 TATCCTAACTAACTGGTCTTCACTCAAGCAGAGTTTCTTCACTCTGGCACTGTGATCAT 1019

QY 1754 GAACTTTAGTAGAGGGGATTTGTGTATTTTATACAAATTTTATACAACTCTTACATTCG 1813
DB 1020 GAACTTTAGTAGAGGGGATTTGTGTATTTTATACAAATTTTATACAACTCTTACATTCG 1079

QY 1814 ATAAAAATTTCTTAAAGAGCAAAACTGCAATTTTATTTCTGCAATCCACATTCOAATCATTA 1873
DB 1080 ATAAAAATTTCTTAAAGAGCAAAACTGCAATTTTATTTCTGCAATCCACATTCOAATCATTA 1139

QY 1874 GAACATAGATATTTATCTATCAAGATATAAATGGTCGAGAGAGACTTTCATCTGTGGATT 1933
DB 1140 GAACATAGATATTTATCTATCAAGATATAAATGGTCGAGAGAGACTTTCATCTGTGGATT 1199

QY 1934 GCGTTGTTTCTTAGGGTTCCTAGCACTGATGCCCTGCACAAGCATGTGATATGTGAATAA 1993
DB 1200 GCGTTGTTTCTTAGGGTTCCTAGCACTGATGCCCTGCACAAGCATGTGATATGTGAATAA 1259

QY 1994 AATGGATCTTCTATAGCTAAATGAGTTCCCTCTGGGAGAGTTCTGGTACTCCATCAC 2053
DB 1260 AATGGATCTTCTATAGCTAAATGAGTTCCCTCTGGGAGAGTTCTGGTACTCCATCAC 1319

QY 2054 AATGCCAGATGGTGTATTTATGGGCTATTTGTGAAGTAAGTGGTAAGATGCTATGAAGTAA 2113
DB 1320 AATGCCAGATGGTGTATTTATGGGCTATTTGTGAAGTAAGTGGTAAGATGCTATGAAGTAA 1379

QY 2114 GTGTGTTTGTTCATCTTATGGAACCTCTGATGCAATGCTTTGTATGGAATAAAT 2173
DB 1380 GTGTGTTTGTTCATCTTATGGAACCTCTGATGCAATGCTTTGTATGGAATAAAT 1439

QY 2174 TTGTGCAATATGATGTCATTCACCTTTGCAATTTGAATTTGCTTTGTTATTTATG 2233
DB 1440 TTGTGCAATATGATGTCATTCACCTTTGCAATTTGAATTTGCTTTGTTATTTATG 1499

QY 2234 TATTTATCTGTGTCAGCTTCTAGTTGCTTCAACCATTTTATTAACATTTTGTACATAT 2293
DB 1500 TATTTATCTGTGTCAGCTTCTAGTTGCTTCAACCATTTTATTAACATTTTGTACATAT 1559

QY 2294 TTACTTGAATAATTTTAAATGGAATTTAAATAAACATTTGATGTTTACAT 2346
DB 1560 TTACTTGAATAATTTTAAATGGAATTTAAATAAACATTTGATGTTTACAT 1612

RESULT 6
AX342061 1626 bp DNA linear PAT 12-JAN-2002
LOCUS
DEFINITION Sequence 24 from Patent EP1166798.
ACCESSION AX342061
VERSION AX342061.1 GI:18151630
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Siemester, G., Haberey, M. and Thierach, K.H.
TITLE Combinations and compositions which interfere with vegf/vegfr and
angiopoietin/tie receptor function and their use
JOURNAL Patent: EP 1166798-A 24 02-JAN-2002;
Schering Aktiengesellschaft (DE)
FEATURES
source
1. 1626
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 68.2%; Score 1599.4; DB 6; Length 1626;
Best Local Similarity 99.9%; Pred. No. 1.6e-276;
Matches 1611; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 734 GGCAATTCCTAGAACTCTAGTAGTATCAGGATATATTTTGGTAAATATATTTGG 793
DB 1 GGCAATTCCTAGAACTCTAGTAGTATCAGGATATATTTTGGTAAATATATTTGG 60

794 TTTATTTGAATACAGACATTGGCTCCCAAAATTTTTCATCTTTGCAAAATAGATGATGACTTTTC 853
Db TTTATTTGAATACAGACATTGGCTCCCAAAATTTTTCATCTTTGCAAAATAGATGATGACTTTTC 120
QY ACTAGAACTTCTCAAATTTGGGAACTTTGCAAAATAGATGATGACTTTTC 913
Db ACTAGAACTTCTCAAATTTGGGAACTTTGCAAAATAGATGATGACTTTTC 180
QY TATCAATTAATGCTATGAGATACATTTGTTTCTCCCTATGCCAAACAGGTGAAACAAACGT 973
Db TATCAATTAATGCTATGAGATACATTTGTTTCTCCCTATGCCAAACAGGTGAAACAAACGT 240
QY AGTTGTTTTTACATGATACATTAATGTTGGCTAATCTGTTGATTTTATAGTATGCAATGTCA 1033
Db AGTTGTTTTTACATGATACATTAATGTTGGCTAATCTGTTGATTTTATAGTATGCAATGTCA 300
QY GAAAAAGGCAAGCAAAATGGCTCTGTTGATGAACTTTGGCAAACTTATGGGCTTTC 1093
Db GAAAAAGGCAAGCAAAATGGCTCTGTTGATGAACTTTGGCAAACTTATGGGCTTTC 360
QY ATTTTCTGACAGACAGGATTTGATCAATATTTGATGAGCTTGGTGAATGATATCAT 1153
Db ATTTTCTGACAGACAGGATTTGATCAATATTTGATGAGCTTGGTGAATGATATCAT 420
QY GGTAGTATGACATGTTAGAAATGGTTTATGTTTATGATCTCGAAATTCATCTCAGGATG 1213
Db GGTAGTATGACATGTTAGAAATGGTTTATGTTTATGATCTCGAAATTCATCTCAGGATG 480
QY AATCTTTTATGCTTTTATTTGTAAGCATATCTGAAATTTTACTTTTATAAAGATGTTTATG 1273
Db AATCTTTTATGCTTTTATTTGTAAGCATATCTGAAATTTTACTTTTATAAAGATGTTTATG 540
QY AAGCTTTTCTCAAAATTTGGCTAGGAATGTTTATCTTCTTCTGATTTGCTGCAAGGCT 1333
Db AAGCTTTTCTCAAAATTTGGCTAGGAATGTTTATCTTCTTCTGATTTGCTGCAAGGCT 600
QY AGAAAAATTAATGTTGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTT 1393
Db AGAAAAATTAATGTTGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTT 660
QY ATGTTATTAATTAATTTTATTTCTGTCAGACAGCATTTTATTTTTCGCAAGTGGAGT 1453
Db ATGTTATTAATTAATTTTATTTCTGTCAGACAGCATTTTATTTTTCGCAAGTGGAGT 720
QY CCATTTAGCCAGTGGGAAAGTCTTGGAACTCAGGTTACCTTGAAGATATGCTGGCAG 1513
Db CCATTTAGCCAGTGGGAAAGTCTTGGAACTCAGGTTACCTTGAAGATATGCTGGCAG 780
QY CCATCTTTGATCTGCTTAAACTGTAATTTATGACACAGCATTAATCCCTTAACTTGGGA 1573
Db CCATCTTTGATCTGCTTAAACTGTAATTTATGACACAGCATTAATCCCTTAACTTGGGA 840
QY TCTGGAATGCAATTTAGTTATGATGCTTGTACCATTTCCAGAAATTTCCAGGGGCAATCGTGGTT 1633
Db TCTGGAATGCAATTTAGTTATG- CCTTGTACCAATTTCCAGAAATTTCCAGGGGCAATCGTGGTT 899
QY TGGTCTAGTGTGATGAAACACAAAGAACAGAGATCCAGCTGAAAAAGAGATGATCTCAA 1693
Db TGGTCTAGTGTGATGAAACACAAAGAACAGAGATCCAGCTGAAAAAGAGATGATCTCAA 959
QY TATCTTAATGCTGCTCTCAATCTCAGACAGAGTTTCTTCACTCTGSCATCTGTGATCAT 1753
Db TATCTTAATGCTGCTCTCAATCTCAGACAGAGTTTCTTCACTCTGSCATCTGTGATCAT 1019
QY GAAACTTAGTAGAGGGGATTTGTGTATTTTATACAAATTTTATACAAATTTTATACAAATTT 1813
Db GAAACTTAGTAGAGGGGATTTGTGTATTTTATACAAATTTTATACAAATTTTATACAAATTT 1079
QY ATAAATTTCTTAAGAGCAAACTGCAATTTTCTGCTCCCAATTTCCAAATCATATTA 1873
Db ATAAATTTCTTAAGAGCAAACTGCAATTTTCTGCTCCCAATTTCCAAATCATATTA 1139

QY 1874 GAACTAAGATATTTATCTATGAGATATAAATGTTGTCAGAGAGACTTTTCATCTGTGGATT 1933
Db 1140 GAACTAAGATATTTATCTATGAGATATAAATGTTGTCAGAGAGACTTTTCATCTGTGGATT 1199
QY 1934 GGGTTGTTTCTTAGGGTTTCTAGCACTGATCCCTGCAAGCATGTGATATGTGAATAA 1993
Db 1200 GGGTTGTTTCTTAGGGTTTCTAGCACTGATCCCTGCAAGCATGTGATATGTGAATAA 1259
QY 1994 AATGATTTCTTCTATGCTAATGATTTCCCTTGGGGAGAGTTCTGTGCTACTGCAATCAC 2053
Db 1260 AATGATTTCTTCTATGCTAATGATTTCCCTTGGGGAGAGTTCTGTGCTACTGCAATCAC 1319
QY 2054 AATGCCAGATGTTTATGCTAATGATTTGCTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGT 2113
Db 1320 AATGCCAGATGTTTATGCTAATGATTTGCTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGT 1379
QY 2114 GTGTGTTTGTCTTCTATGCTAATGATTTCCCTTGGGGAGAGTTCTGTGCTACTGCAATCAC 2173
Db 1380 GTGTGTTTGTCTTCTATGCTAATGATTTCCCTTGGGGAGAGTTCTGTGCTACTGCAATCAC 1439
QY 2174 TTGGTGCATATGATGCTCAATGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 2233
Db 1440 TTGGTGCATATGATGCTCAATGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1499
QY 2234 TATTATACCTGTCAGCTTCTAGTTGCTTCAACCAATTTTATAACCAATTTTGTACATATT 2293
Db 1500 TATTATACCTGTCAGCTTCTAGTTGCTTCAACCAATTTTATAACCAATTTTGTACATATT 1559
QY 2294 TTACTTGAATAATTTTAAATGGAATTTTAAATGGAATTTTAAATGGAATTTTAAATGGAATTT 2346
Db 1560 TTACTTGAATAATTTTAAATGGAATTTTAAATGGAATTTTAAATGGAATTTTAAATGGAATTT 1612

RESULT 7
AX342121 1626 bp DNA linear PAT 12-JAN-2002
LOCUS AX342121
DEFINITION Sequence 24 from Patent EP1166799.
ACCESSION AX342121
VERSION AX342121.1 GI:18151689
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Siemester, G., Haberey, M. and Thierach, K.H.
TITLE Combinations and compositions which interfere with vegf/vegf and
angiopoietin/tie receptor function and their use (ii)
JOURNAL Patent: EP 1166799-A 24 02-JAN-2002;
SCHERING AKTIEGENSCHAFT (DE)
FEATURES
Location/Qualifiers
source 1..1626
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 68.2%; Score 1599.4; DB 6; Length 1626;
Best Local Similarity 99.9%; Pred. No. 1.6e-276;
Matches 1611; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 734 GGCAATTTCTAGATCTATAGTAGTATCAGGATATATTTTCTTTTAAATATATTTTGG 793
Db 1 GGCAATTTCTAGATCTATAGTAGTATCAGGATATATTTTCTTTTAAATATATTTTGG 60
QY 794 TTTATTTGAAATCAGACATTTGCTCCAAATTTTTCATCTTTGCAACAATAGTATGACTTTTC 853
Db 61 TTTATTTGAAATCAGACATTTGCTCCAAATTTTTCATCTTTGCAACAATAGTATGACTTTTC 120
QY 854 ACTAGAACTTCTCAACATTTGGGAATTTTGGCAATATGAGCATCATATGTTTAAAGGCTG 913
Db 121 ACTAGAACTTCTCAACATTTGGGAATTTTGGCAATATGAGCATCATATGTTTAAAGGCTG 180

QY 1034 GAAAAGCGACAGAAATGGGCTCTTGTAAGTAACTTTCGGCAAACTTATGGGTCTTC 1093
DB 301 GAAAAGCGACAGAAATGGGCTCTTGTAAGTAACTTTCGGCAAACTTATGGGTCTTC 360
QY 1094 ATTTTCTGACACACAGGATTTGACTCAATATTTGTAGAGCTTGGTGAATGGAATTACAT 1153
DB 361 ATTTTCTGACACACAGGATTTGACTCAATATTTGTAGAGCTTGGTGAATGGAATTACAT 420
QY 1154 GGTAGTATGACACTGGTGAAGATGTTTATGTTATGATCAGATCAGAAATTCATCTCAGGATG 1213
DB 421 GGTAGTATGACACTGGTGAAGATGTTTATGTTATGATCAGATCAGAAATTCATCTCAGGATG 480
QY 1214 AATCTTTTATGCTTTTATGTTAGTATGATATCTGAATTTTATCTTTATAAGATGTTTATG 1273
DB 481 AATCTTTTATGCTTTTATGTTAGTATGATATCTGAATTTTATCTTTATAAGATGTTTATG 540
QY 1274 AAGCTTTTGTCTAAATAATTTGGCTTAGGAATGGTAACCTTCAATTTTCAGTTGCCAAGGGGT 1333
DB 541 AAGCTTTTGTCTAAATAATTTGGCTTAGGAATGGTAACCTTCAATTTTCAGTTGCCAAGGGGT 600
QY 1334 AGAAAAATAATATGTTGTTATGTTATGTTTAAACATATTTATGAGTACTATCTATGA 1393
DB 601 AGAAAAATAATATGTTGTTATGTTTAAACATATTTATGAGTACTATCTATGA 660
QY 1394 ATGATTTTAAATPATTTTTCATATTTCTGTGACAGCAATTTATAATTTGCAACAAGTGGAGT 1453
DB 661 ATGATTTTAAATPATTTTTCATATTTCTGTGACAGCAATTTATAATTTGCAACAAGTGGAGT 720
QY 1454 CAATTAGCCAGTGGGAAGCTTTGGAACTCAGGTTACCCCTTGAAGATATGCTGGCAG 1513
DB 721 CAATTAGCCAGTGGGAAGCTTTGGAACTCAGGTTACCCCTTGAAGATATGCTGGCAG 780
QY 1514 CCATCTCTTTGATCTGTGCTTAAATCTGTAATTTATAGACCAGCTAAATCCCTTAATCTTGG 1573
DB 781 CCATCTCTTTGATCTGTGCTTAAATCTGTAATTTATAGACCAGCTAAATCCCTTAATCTTGG 840
QY 1574 TCTGGAATGCATATGTTATGACTTGTACCTTCCAGAAATTTCAAGGGCACTGTTGGT 1633
DB 841 TCTGGAATGCATATGTTATG-CTTGTACCAATTCAGGAAATTTCAAGGGCACTGTTGGT 899
QY 1634 TGGTCTAGTGTGTAAGAAACAAAGAACACAGAGATCCAGCTGAAAAGAGTATCCTCAA 1693
DB 900 TGGTCTAGTGTGTAAGAAACAAAGAACACAGAGATCCAGCTGAAAAGAGTATCCTCAA 959
QY 1694 TATCTTAATCTAGTCTCAACTCAAGCAGAGTTTCTTCACTCTGGCAGCTGATCAT 1753
DB 960 TATCTTAATCTAGTCTCAACTCAAGCAGAGTTTCTTCACTCTGGCAGCTGATCAT 1019
QY 1754 GAAACTTAGTAGAGGGGATTTGTGTATTTTATACAAATTTTAAATCAATGTCTTACATTG 1813
DB 1020 GAAACTTAGTAGAGGGGATTTGTGTATTTTATACAAATTTTAAATCAATGTCTTACATTG 1079
QY 1814 ATAAATTTCTTAAAGAGCAAACTGCATTTTATTTCTGCAATCCAGATTCGAATCATATTA 1873
DB 1080 ATAAATTTCTTAAAGAGCAAACTGCATTTTATTTCTGCAATCCAGATTCGAATCATATTA 1139
QY 1874 GAACTAAGATATTTATCTATGAGATATAAATGGTGCAGAGACTTTTCATCTGGGAT 1933
DB 1140 GAACTAAGATATTTATCTATGAGATATAAATGGTGCAGAGACTTTTCATCTGGGAT 1199
QY 1934 GGTCTTTTCTTAGGGTCTTAGCACTGATGCTGCAAGCATGTGATATGTAATGAAATAA 1993
DB 1200 GGTCTTTTCTTAGGGTCTTAGCACTGATGCTGCAAGCATGTGATATGTAATGAAATAA 1259
QY 1994 AATGGAATCTTCTATAGCTAAATGAGTTCCTCTGGGAGAGTTCCTGCTAGCTGCAATC 2053
DB 1260 AATGGAATCTTCTATAGCTAAATGAGTTCCTCTGGGAGAGTTCCTGCTAGCTGCAATC 1319
QY 2054 AATGCCAGATGTTGTTTATGGGCTATTTGTGTAAGTAAGTGTGTAAGATCTATGAATAA 2113
DB 1320 AATGCCAGATGTTGTTTATGGGCTATTTGTGTAAGTAAGTGTGTAAGATCTATGAATAA 1379
QY 2114 GTGTGTTTGTTCATCTTATGGAATACTCTGTGATGCAATGCTTTTGTATGGAATAAT 2173

DB 1380 GTGTGTTTGTTCATCTTATGGAATACTCTTGTGATGCTGCTTTTGTATGGAATAAT 1439
QY 2174 TTGGTCAATATGATGATCAATTTCAACTTTTCATTTGAATTTGGTGTGTTATATG 2233
DB 1440 TTGGTCAATATGATGATCAATTTCAACTTTTCATTTGAATTTGGTGTGTTATATG 1499
QY 2234 TATTATACCTGTCCAGCTTCTAGTTGCTTCAACCAATTTTATACCAATTTTGTACATAT 2293
DB 1500 TATTATACCTGTCCAGCTTCTAGTTGCTTCAACCAATTTTATACCAATTTTGTACATAT 1559
QY 2294 TTACTTTGAAATAATTTTAAATGGAATTTTAAATAAACAATTTGATAGTTTACAT 2346
DB 1560 TTACTTTGAAATAATTTTAAATGGAATTTTAAATAAACAATTTGATAGTTTACAT 1612

RESULT 9
AX351446/c
LOCUS AX351446 586 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 193 from Patent WO0196390.
ACCESSION AX351446
VERSION AX351446.1 GI:18616793
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Jiang, Y., Hepler, W. T., Clapper, J. D., Wang, A. and Secrist, H.
TITLE Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL Patent: WO 0196390-A 193 20-DEC-2001;
CORIXA CORPORATION (US)
FEATURES
Location/Qualifiers
source 1..586
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 24.8%; Score 582.8; DB 6; Length 586;
Best Local Similarity 99.7%; Pred. No. 2e-94;
Matches 584; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 476 AACTTCTTATTCCTCTGAAAGAAAGCAATGTAATCACTATGATTTATTCACAAA 535
DB 586 AACTTCTTATTCCTCTGAAAGAAAGCAATGTAATCACTATGATTTATTCGCAAA 527
QY 536 CAACCAAGAAATTCCTCAACAAATTTTAAAGTAATCTGATCCTCTCTTGGAGAAAATTTGTAC 595
DB 526 CAACCAAGAAATTCCTCAACAAATTTTAAAGTAATCTGATCCTCTCTTGGAGAAAATTTGTAC 467
QY 596 CTAATAGTTTTCTTATGATGATGTTATTTACTACTGCTATTAATCAATTTCTATAAATTT 655
DB 466 CTAATAGTTTTCTTATGATGATGTTATTTACTACTGCTATTAATCAATTTCTATAAATTT 407
QY 656 CCTACTTAAAGTCTTAAGAACTGGGTTCTTCTTTGATGTTTATTCATGTTTTCAGAAAGGAAA 715
DB 406 CCTACTTAAAGTCTTAAGAACTGGGTTCTTCTTTGATGTTTATTCATGTTTTCAGAAAGGAAA 347
QY 716 CAACACTTTACTCTTTTAGGCAATTCCTAGAAATCTATAGTAGTATCAGGATATATTTTG 775
DB 346 CAACACTTTACTCTTTTAGGCAATTCCTAGAAATCTATAGTAGTATCAGGATATATTTTG 287
QY 776 CTTTAAATAATATTTTGGTTATTTTGAATACAGACATTTGGCTCCAAATTTTCATCTTTGC 835
DB 286 CTTTAAATAATATTTTGGTTATTTTGAATACAGACATTTGGCTCCAAATTTTCATCTTTGC 227
QY 836 ACAATAGTATGACTTTTCACTAGAACTTCTCAACATTTGGGAACTTTGCAAAATATGAGCA 895
DB 226 ACAATAGTATGACTTTTCACTAGAACTTCTCAACATTTGGGAACTTTGCAAAATATGAGCA 167
QY 896 TCATATGTTTAAAGCTGTATCATTTTAACTATGATGATCAATTTGTTTCTCCCTATGCC 955

```
Db 166 TCATATGTTGTAAGGCTGTATCATTTAAATGCTATGAGATACATGTTGTTCTCCCTATGCC 107
Qy 956 AAACAGGTGAACAAACGCTAGTCTTTTACTGATCTAAATGTTGGCTACCTGTGATTT 1015
Db 106 AAACAGGTGAACAAACGCTAGTCTTTTACTGATCTAAATGTTGGCTACCTGTGATTT 47
Qy 1016 TATAGTATGCACATGTCAGAAAGGCAAGCAAAATGGCTCTTGT 1061
Db 46 TATAGTATGCACATGTCAGAAAGGCAAGCAAAATGGCTCTTGT 1

RESULT 10
LOCUS AR273526 675 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1269 from patent US 6504010.
ACCESSION AR273526
VERSION AR273526.1 GI:29705411
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 675)
AUTHORS Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,
Carter,D., Retter,M.W., Mannion,J. and Fan,L.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: US 6504010-A 1269 07-JAN-2003;
FEATURES Location/Qualifiers
source 1..675
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 24.8%; Score 581.4; DB 6; Length 675;
Best Local Similarity 97.3%; Pred. No. 3.4e-94;
Matches 622; Conservative 0; Mismatches 13; Indels 4; Gaps 3;
Qy 1673 CTGAAAAGAGTGATCTCAATATCCTAACTAACTGGTCTCACTCAAGCAGAGTTTCT 1732
Db 1 CTGAAAAGAGTGATCTCAATATCCTAACTAACTGGTCTCACTCAAGCAGAGTTTCT 60
Qy 1733 TCACCTCTGGCAGCTGATCATGAACTTAGTAGAGGGGATGTTGTATTTATACAAAT 1792
Db 61 TCACCTCTGGCAGCTGATCATGAACTTAGTAGAGGGGATGTTGTATTTATACAAAT 120
Qy 1793 TTAATACAATGTCTTACATGATAAAATCTTAAAGAGCAAAACTGCATTTTATTTCTGC 1852
Db 121 TTAATACAATGTCTTACATGATAAAATCTTAAAGAGCAAAACTGCATTTTATTTCTGC 180
Qy 1853 ATCCACATTCCAATCATATTAGAACTAAGATATTTATCTATGAAGATATAAATGGTGCAG 1912
Db 181 ATCCACATTCCAATCATATTAGAACTAAGATATTTATCTATGAAGATATAAATGGTGCAG 240
Qy 1913 AGAGACTTTTCATCTGTGGATGGGTTGTTCTTAGGGTTCCTAGCACTGATGCTGCACA 1972
Db 241 AGAGACTTTTCATCTGTGGATGGGTTGTTCTTAGGGTTCCTAGCACTGATGCTGCACA 300
Qy 1973 AGCATGTGATATGTAATAAATGGAATCTTCTATAGCTAAATGAGTCCCTCTGGGGA 2032
Db 301 AGCATGTGATATGTAATAAATGGAATCTTCTATAGCTAAATGAGTCCCTCTGGGGA 360
Qy 2033 GAGTCTCTGGTACTGCAATCAATGCCAGATGGTGTATTTATGGGCTATTTGTGTAAAGTAAG 2092
Db 361 GAGTCTCTGGTACTGCAATCAATGCCAGATGGTGTATTTATGGGCTATTTGTGTAAAGTAAG 420
Qy 2093 TGGTAAAGATGCTATGAAGTAAGTGTGTTTGTGTTTTCATCTTATGAAACTCTTGTATGATG 2152
Db 421 TGGTAAAGATGCTATGAAGTAAGTGTGTTTGTGTTTTCATCTTATGAAACTCTTGTATGATG 480
Qy 2153 TGCCTTTGTATGGAATAAAATTTGGTCAATATGATGTCATTCAACTTTGTCATTTGAATTG 2212
Db 481 TGCCTTTGTATGGAATAAAATTTGGTCAATATGATGTCATTCAACTTTGTCATTTGAATTG 540
```

```
Qy 2213 -AAATTTGGTGTATTTATATGATTTATA-CTGTCAACGCTTCTAGTTGCTTCAACCAATT 2270
Db 541 AAAATTTGGGTGATTTATATGATTTATATACCTGTCAACGCTTCTAGTTGCTTCAACCAATT 600
Qy 2271 TTAATAACCAATTTTGTACATATTTTACTTGAATAATTTT 2309
Db 601 T--ATACATTTTGNACATATTTTACTTGNAAATATTT 637

RESULT 11
LOCUS AR277107 675 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1269 from patent US 6509448.
ACCESSION AR277107
VERSION AR277107.1 GI:29710754
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 675)
AUTHORS Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S.,
Carter,D., Retter,M.W., Mannion,J., Fan,L. and Wang,A.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: US 6509448-A 1269 21-JAN-2003;
FEATURES Location/Qualifiers
source 1..675
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 24.8%; Score 581.4; DB 6; Length 675;
Best Local Similarity 97.3%; Pred. No. 3.4e-94;
Matches 622; Conservative 0; Mismatches 13; Indels 4; Gaps 3;
Qy 1673 CTGAAAAGAGTGATCTCAATATCCTAACTAACTGGTCTCACTCAAGCAGAGTTTCT 1732
Db 1 CTGAAAAGAGTGATCTCAATATCCTAACTAACTGGTCTCACTCAAGCAGAGTTTCT 60
Qy 1733 TCACCTCTGGCAGCTGATCATGAACTTAGTAGAGGGGATGTTGTATTTATACAAAT 1792
Db 61 TCACCTCTGGCAGCTGATCATGAACTTAGTAGAGGGGATGTTGTATTTATACAAAT 120
Qy 1793 TTAATAACAATGTCTTACATGATAAAATCTTAAAGAGCAAAACTGCATTTTATTTCTGC 1852
Db 121 TTAATAACAATGTCTTACATGATAAAATCTTAAAGAGCAAAACTGCATTTTATTTCTGC 180
Qy 1853 ATCCACATTCCAATCATATTAGAACTAAGATATTTATCTATGAAGATATAAATGGTGCAG 1912
Db 181 ATCCACATTCCAATCATATTAGAACTAAGATATTTATCTATGAAGATATAAATGGTGCAG 240
Qy 1913 AGAGACTTTTCATCTGTGGATGGGTTGTTCTTAGGGTTCCTAGCACTGATGCTGCACA 1972
Db 241 AGAGACTTTTCATCTGTGGATGGGTTGTTCTTAGGGTTCCTAGCACTGATGCTGCACA 300
Qy 1973 AGCATGTGATATGTAATAAATGGAATCTTCTATAGCTAAATGAGTCCCTCTGGGGA 2032
Db 301 AGCATGTGATATGTAATAAATGGAATCTTCTATAGCTAAATGAGTCCCTCTGGGGA 360
Qy 2033 GAGTCTCTGGTACTGCAATCAATGCCAGATGGTGTATTTATGGGCTATTTGTGTAAAGTAAG 2092
Db 361 GAGTCTCTGGTACTGCAATCAATGCCAGATGGTGTATTTATGGGCTATTTGTGTAAAGTAAG 420
Qy 2093 TGGTAAAGATGCTATGAAGTAAGTGTGTTTGTGTTTTCATCTTATGAAACTCTTGTATGATG 2152
Db 421 TGGTAAAGATGCTATGAAGTAAGTGTGTTTGTGTTTTCATCTTATGAAACTCTTGTATGATG 480
Qy 2153 TGCCTTTGTATGGAATAAAATTTGGTCAATATGATGTCATTCAACTTTGTCATTTGAATTG 2212
Db 481 TGCCTTTGTATGGAATAAAATTTGGTCAATATGATGTCATTCAACTTTGTCATTTGAATTG 540
Qy 2213 -AAATTTGGTGTATTTATATGATTTATA-CTGTCAACGCTTCTAGTTGCTTCAACCAATT 2270
```

||||| 541 AAATTTGGGTGATTTATATGATTTATACCGTGCAGCTTCTAGTTGCTTCAACCAAT 600
QY 2271 TTATAACCAATTTTGTACATATTTTACCTGAAATATTT 2309
Db 601 T--ATACCAATTTTGNACATATTTTACTTGNAAATATTT 637

RESULT 12
AX368559 675 bp DNA linear PAT 16-FEB-2002
LOCUS AX368559
DEFINITION Sequence 1269 from Patent WO0204514.
ACCESSION AX368559
VERSION AX368559.1 GI:18856632
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W.,
AUTHORS Marnerakis, M., Carter, D., Fanger, G.R., Vedvick, T.S., Bangur, C.S.,
McNabb, A., Fanger, N., Switzer, A., McNeill, P.D. and Clapper, J.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer
JOURNAL Patent: WO 0204514-A 1269 17-JAN-2002;
CORIXA CORPORATION (US)
FEATURES
source
1. 675
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 24.8%; Score 581.4; DB 6; Length 675;
Best Local Similarity 97.3%; Pred. No. 3.4e-94;
Matches 622; Conservative 0; Mismatches 13; Indels 4; Gaps 3;

QY 1673 CTGAAAAGAGTGATCCTCAATATCTAATCTAACTGGTCTCAAGCAGAGTTCT 1732
Db 1 CTGAAAAGAGTGATCCTCAATATCTAATCTAACTGGTCTCAAGCAGAGTTCT 60

QY 1733 TCACCTCGCATGATCATGAATCTAGTAGAGGGGATGTGTATATTTATACAAT 1792
Db 61 TCACCTCGCATGATCATGAATCTAGTAGAGGGGATGTGTATATTTATACAAT 120

QY 1793 TTAATACATGCTTACATGTATTAATTTCTTAAAGACAAACGCAATTTATTTCTGC 1852
Db 121 TTAATACATGCTTACATGTATTAATTTCTTAAAGACAAACGCAATTTATTTCTGC 180

QY 1853 ATCCACATTCACATATAGAACTAAGATATTTATCTATAGAAATATAATGGTGAG 1912
Db 181 ATCCACATTCACATATAGAACTAAGATATTTATCTATAGAAATATAATGGTGAG 240

QY 1913 AGAGATTTCACTGTGGATCGGTGTTCTTAGGTTCTTAGCAGTGCCTGCGACA 1972
Db 241 AGAGATTTCACTGTGGATCGGTGTTCTTAGGTTCTTAGCAGTGCCTGCGACA 300

QY 1973 AGCATGTGATATGTGAAATATAATGGAATCTTCTATAGCTAAATGAGTTCCCTCTGGGGA 2032
Db 301 AGCATGTGATATGTGAAATATAATGGAATCTTCTATAGCTAAATGAGTTCCCTCTGGGGA 360

QY 2033 GAGTTCTGTAGTGCATCAATGAGATCGGTGTTCTTAGGTTCTTAGGTTCTTAGGTTCTAG 2092
Db 361 GAGTTCTGTAGTGCATCAATGAGATCGGTGTTCTTAGGTTCTTAGGTTCTTAGGTTCTAG 420

QY 2093 TGGTAAGATGCTATGAAAGTGTGTTTGTGTTTCTATCTTATGGAACCTTTGATGCGATG 2152
Db 421 TGGTAAGATGCTATGAAAGTGTGTTTGTGTTTCTATCTTATGGAACCTTTGATGCGATG 480

QY 2153 TGCCTTTGTATGGAATAAATTTTGGTGCAATATGATGATGATGATGATGATGATGATGATG 2212
Db 481 TGCCTTTGTATGGAATAAATTTTGGTGCAATATGATGATGATGATGATGATGATGATGATG 540

QY 2213 -AATTTGGTGTATTTATATGATTTATATA-CCTGTCAACGCTTCTAGTTGCTTCAACCAAT 2270
Db 541 AAATTTGGGTGATTTATATGATTTATATACCGTGCAGCTTCTAGTTGCTTCAACCAAT 600

QY 2271 TTATAACCAATTTTGTACATATTTTACCTGAAATATTT 2309
Db 601 T--ATACCAATTTTGNACATATTTTACTTGNAAATATTT 637

RESULT 13
AX341176 586 bp DNA linear PAT 10-JAN-2002
LOCUS AX341176
DEFINITION Sequence 1423 from Patent WO0196388.
ACCESSION AX341176
VERSION AX341176.1 GI:18137158
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Jiang, Y., Harlocker, S.L. and Secrist, H.
AUTHORS Compositions and methods for the therapy and diagnosis of colon
TITLE cancer
JOURNAL Patent: WO 0196388-A 1423 20-DEC-2001;
CORIXA CORPORATION (US)
FEATURES
source
1. 586
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 24.2%; Score 568; DB 6; Length 586;
Best Local Similarity 99.0%; Pred. No. 8.7e-92;
Matches 580; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 464 ACTACCCCAATTAACCTTCTTATTGCTCTGAAGAAGAAAGCAATGTAATCACTATGA 523
Db 1 ACTACCCCAATTAACCTTCTTATTGCTCTGAAGAAGAAAGCAATGTAATCACTATGA 60

QY 524 TTATTGCAAAACACAGAAATCTCCAACATTTTAAAGTAATCTGATCTCTCTTGGGA 593
Db 61 TTATTGCGCAACACAGAAATCTCCAACATTTTAAAGTAATCTGATCTCTCTTGGGA 120

QY 584 GAAATTTGTACCTAATAGTTTCTTATGATGATTTTACTCTAGTATTAATCAATCAAT 643
Db 121 GAAATTTGTACCTAATAGTTTCTTATGATGATTTTACTCTAGTATTAATCAATCAAT 180

QY 644 TTCTATAAATTTCTTAACTTAAGCTTTAAGAACTGGGTTCTTCTTTGATGTTATTCATGT 703
Db 181 TTCTATAAATTTCTTAACTTAAGCTTTAAGAACTGGGTTCTTCTTTGATGTTATTCATGT 240

QY 704 TCAGAAAGGAAACACACATTTTACTCTTTAGGACAATTTCTAGAACTCTATAGTAGTATCA 763
Db 241 TCAGAAAGGAAACACACATTTTACTCTTTAGGACAATTTCTAGAACTCTATAGTAGTATCA 300

QY 764 GGATATATTTTGTGTTTAAATATATTTTGGTATTTTGAATACAGACATTTGGCTCCAAAT 823
Db 301 GGATATATTTTGTGTTTAAATATATTTTGGTATTTTGAATACAGACATTTGGCTCCAAAT 360

QY 824 TTTTCATCTTTGCAATAGTATGACTTTTCTCAGTAACTTTCTCAACATTTGGGAACCTTTG 883
Db 361 TTTTCATCTTTGCAATAGTATGACTTTTCTCAGTAACTTTCTCAGTAACTTTGGGAACCTTTG 420

QY 884 CAAATATGAGCATCATATGTTTAAAGCTG-TATCAATTTAAATGCTATAGATACATGTT 942
Db 421 CAAATATGAGCATCATATGTTTAAAGCTGTTATCATTTAAATGCTATAGATACATGTT 480

QY 943 TTCTCCCTATGCCAAACAGGTGAACAACTGTTCTTTTACTGATCACTAATGTTGG 1002
Db 481 TTCTCCCTATGCCAAACAGGTGAACAACTGTTCTTTTACTGATCACTAATGTTGG 540

QY	1003	CTACCTGTGATTTTATAGTATGCATGTCAGAAAGGCGAGACA	1048
Db	541	CTACCTGNGATTTTATAGTATGCATGTGAGAAAGGCGAGDNA	596
RESULT 14			
AR407382			
LOCUS	AR407382	665 bp	DNA
DEFINITION	Sequence 1269 from patent US 6630574.		linear
ACCESSION	AR407382		
VERSION	AR407382.1	GI:40157193	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 665)		
AUTHORS	Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S., Carter,D., Ketter,M.W., Mannion,J. and Fan,L.		
TITLE	Compositions and methods for the therapy and diagnosis of lung cancer		
JOURNAL	Patent: US 6630574-A	1269 07-OCT-2003;	
FEATURES	Location/Qualifiers		
source	1..665		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
	Query Match	23.5%;	Score 551.4; DB 6; Length 665;
	Best Local Similarity	95.8%;	Pred. No. 8e-89;
	Matches	612;	Conservative 0; Mismatches 13; Indels 14; Gaps 4;
QY	1673	CTGAAAAGAGTGTATCCTCAATATCCTAACTGCTCCTCAACTCAAGCAGAGTTTCT	1732
Db	1	CTGAAAAGAGTGTATCCTCAATATCCTAACTGCTCCTCAACTCAAGCAGAGTTTCT	60
QY	1733	TCACCTCGGCACCTGTGATCATGAAACTTAGTAGAGGGAGTGTGTATTTATACAAAT	1792
Db	61	TCACCTCGGCACCTGTGATCATGAAACTTAGTAGAGGGAGTGTGTATTTATACA	117
QY	1793	TTAATACAATGCTCTTACATGATGATAAATCTTAAAGAGCAAACTGCATTTATTCGC	1852
Db	118	-----AATGCTCTTACATGATGATAAATCTTAAAGAGCAAACTGCATTTATTCGC	170
QY	1853	ATCCACATTCCAATCATATTAGAACTAAGATATTATCTATGAAGATATAAATGGTCGAG	1912
Db	171	ATCCACATTCCAATCATATTAGAACTAAGATATTATCTATGAAGATATAAATGGTCGAG	230
QY	1913	AGACATTTCACTGTGGATTCGTTCTTTAGGGTTCCCTAGCAGTGCTCGACA	1972
Db	231	AGACATTTCACTGTGGATTCGTTCTTTAGGGTTCCCTAGCAGTGCTCGACA	290
QY	1973	AGCATGTGATATGGAATAAATAGGATCTTCTATAGCTAAATAGTTCCTCTCGGGA	2032
Db	291	AGCATGTGATATGGAATAAATAGGATCTTCTATAGCTAAATAGTTCCTCTCGGGA	350
QY	2033	GAGTTCTGGTACTGCAATCAAAATGCCAGATGGTGTATGGGCTATTGTGTAAGTAAG	2092
Db	351	GAGTTCTGGTACTGCAATCAAAATGCCAGATGGTGTATGGGCTATTGTGTAAGTAAG	410
QY	2093	TGGTAAGATGCTATGAAGTAAGTCGTGTGTGTGTTCATCTTATGAAACTCTTGTATGCATG	2152
Db	411	TGGTAAGATGCTATGAAGTAAGTCGTGTGTGTGTTCATCTTATGAAACTCTTGTATGCATG	470
QY	2153	TGCTTTTGTATGGAATAAATTTGGTGCAATATGATGTCATTCACATTTTGCATTGAATTG	2212
Db	471	TGCTTTTGTATGGAATAAATTTGGTGCAATATGATGTCATTCACATTTTGCATTGAATTG	530
QY	2213	-AATTTTGGTGTATTTATATGTAATTATA-CCTGTCACGGCTTCAGTGTGCTTCAACCATT	2270
Db	531	AAATTTTGGGFGAATTTATATGTAATTATAOCCTCTCACGCTTCAGTGTGCTTCAACCATT	590
QY	2271	TTATAACCAATTTTGTACATATTTTACTTGGAAATATTTT	2309

```

Db 1694 ATGGTGTATTATGGGCTATTTGTGTAAGTAAAGATGCTATGAAGTAAGTGTGTTT 1753
Qy 2122 GTTTTCATCTTATGGAACCTCTTCATGCAATGCTTTTGTATGGAATAAAATTTGGTGCA 2181
Db 1754 GTTTTCATCTTATGGAACCTCTTCATGCAATGCTTTTGTATGGAATAAAATTTGGTGCA 1813
Qy 2182 ATATGATGTCATTTCAACTTTGCAATGCAATTTTGGTTGTATTTATATGATTTATAC 2241
Db 1814 ATATGATGTCATTTCAACTTTGCAATGCAATTTTGGTTGTATTTATATGATTTATAC 1873
Qy 2242 CTGTCACGCTTCTAGTTGCTTCAACCAATTTTATTAACCAATTTTGTACATATTTTACTTGA 2301
Db 1874 CTGTCACGCTTCTAGTTGCTTCAACCAATTTTATTAACCAATTTTGTACATATTTTACTTGA 1933
Qy 2302 AAATATTTTAAATGGAATTTTAAATTAACCAATTTGATAGTTTACAT 2346
Db 1934 AAATATTTTAAATGGAATTTTAAATTAACCAATTTGATAGTTTACAT 1978

```

Search completed: April 25, 2004, 18:58:41
Job time : 6182.51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 25, 2004, 07:00:48 ; Search time 4507.01 Seconds
(without alignments)

16458.286 Million cell updates/sec

Title: US-10-051-835-18

Perfect score: 2484

Sequence: 1 gttattctgtattataac.....ataaatttttataaaactc 2484

Scoring table: IDENTITY.NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST:
1: em_estba:*
2: em_esthum:*
3: em_estini:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_esti:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	625.4	25.2	627	13 BX470854	BX470854 DKFZp686B
2	617.6	24.9	778	12 BG676811	BG676811 602623406
3	523.8	21.1	607	9 A180763	A180763 zp43f01.r
4	513	20.7	513	9 AA480136	AA480136 zv42f09.r

C	5	510	20.5	510	10	BE049439
C	6	504	20.3	639	14	N90350
C	7	472.8	19.0	504	9	AA031928
C	8	453	18.2	608	10	AW381442
C	9	450	18.1	451	9	AI275105
C	10	442.6	17.8	1201	9	AL541045
C	11	440.8	17.7	452	10	AW130518
C	12	436	17.6	436	9	AI290106
C	13	434.8	17.5	462	14	N33172
C	14	433.4	17.4	763	14	CA411843
C	15	431	17.4	512	9	AA032029
C	16	430.4	17.3	761	14	CA428666
C	17	430	17.3	430	9	AA476356
C	18	427.8	17.2	512	9	AW751313
C	19	427	17.2	1089	9	AL577084
C	20	424.8	17.1	764	13	BO772889
C	21	424.4	17.1	765	14	CA423905
C	22	413.4	16.6	415	10	AW613660
C	23	403.8	16.3	978	13	EX369387
C	24	389.4	15.7	489	10	BF375154
C	25	383.6	15.4	507	14	N90723
C	26	382.8	15.4	615	10	BE826206
C	27	368	14.8	482	9	AA404636
C	28	362.4	14.6	487	12	BM987655
C	29	350.4	14.1	691	10	AW118825
C	30	345	13.9	351	9	AM028857
C	31	341.6	13.8	714	9	AI924840
C	32	328.4	13.2	658	14	CA415309
C	33	324.4	13.1	644	10	AW593610
C	34	324.2	13.1	657	10	AW993106
C	35	317.2	12.8	374	9	AI061108
C	36	313	12.6	424	9	AA233637
C	37	311	12.5	386	10	AW204634
C	38	311	12.5	446	14	CB143712
C	39	311	12.5	483	9	AA775662
C	40	311	12.5	485	9	AI160736
C	41	311	12.5	490	9	AI204255
C	42	311	12.5	496	9	AI351758
C	43	311	12.5	496	10	AW192267
C	44	311	12.5	543	9	AA975121
C	45	311	12.5	565	9	AI565967

ALIGNMENTS

RESULT 1
BX470854
LOCUS
DEFINITION
DKFZp686B06123 r1 686 (synonym: hicc3) Homo sapiens cDNA clone
627 bp mRNA linear EST 04-SEP-2003
DKFZp686B06123 5', mRNA sequence.
ACCESSION
BX470854
VERSION
BX470854.1 GI:31665178
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 627)
AUTHORS
Fahr, A., Lauber, J., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Robo, G., Han, W. and Wiemann, S.
TITLE
EST (Fahr, A., Lauber, J., Mewes, H.W., Weil, B., et al.)
JOURNAL
Unpublished (2003)
COMMENT
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl sequence available.
This clone (DKFZp686B06123) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES		source	
Location/Qualifiers		1..627	
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="DKFZP686B06123"			
/dev_stage="adult"			
/lab_host="DH10B"			
/clone_lib="686 (synonym: hlcc3)"			
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB; cDNA-collection"			
ORIGIN			
Query Match		25.2%; Score 625.4; DB 13; Length 627;	
Best Local Similarity		99.8%; Pred. No. 2.3e-110;	
Matches 626; Conservative 0;		Mismatches 1; Indels 0; Gaps 0;	
QY	19	ACCATAACAGTTCACATAAATAAATTAAGGAATGAATGTTAAATTTGGTT	78
Db	1	ACCATAACAGTTCACATAAATAAATTAAGGAATGAATGTTAAATTTGGTT	60
QY	79	TTATATTTATGTTTACATTTATGCTGTTCAAGATCAAAATCTACAAATATATGATATAGT	138
Db	61	TTATATTTATGTTTACATTTATGCTGTTCAAGATCAAAATCTACAAATATATGATATAGT	120
QY	139	CAAAGAATCTATCTCTCTGCCCCCAATAAATTTCTCCCTCTTCCCAATAGT	198
Db	121	CAAAGAATCTATCTCTCTGCCCCCAATAAATTTCTCCCTCTTCCCAATAGT	180
QY	199	AACCATATAAATTTATTTACTGCTTTTAAATATGTAACAAAGTACATATAAAT	258
Db	181	AACCATATAAATTTATTTACTGCTTTTAAATATGTAACAAAGTACATATAAAT	240
QY	259	TTGCTGCTACTCCCTCTTAGAGAAGTGGTAGAAAATCTATGTTATTTGACTTATCAGAC	318
Db	241	TTGCTGCTACTCCCTCTTAGAGAAGTGGTAGAAAATCTATGTTATTTGACTTATCAGAC	300
QY	319	ATTGTTTAACTGACATGGCATTTTCTGCTACAAATGTTCCAGCAGTAAATCTTTGC	378
Db	301	ATAGTTTAACTGACATGGCATTTTCTGCTACAAATGTTCCAGCAGTAAATCTTTGC	360
QY	379	ATATATCATTTTGCATTTTGTGCTGATATATCAGTGGGACAGATTTCCCAACAGAGAAGTGC	438
Db	361	ATATATCATTTTGCATTTTGTGCTGATATATCAGTGGGACAGATTTCCCAACAGAGAAGTGC	420
QY	439	TAGATGAAGATGAAGTATCATCTGACCTTGAGACCCCTTCCCTCTGCGCTCTGCTGATG	498
Db	421	TAGATGAAGATGAAGTATCATCTGACCTTGAGACCCCTTCCCTCTGCGCTCTGCTGATG	480
QY	499	GTACCAAGCTTTTCTGAACTCTTACTTCCATCTGTGGGACAGATTCCTCTCTTAT	558
Db	481	GTACCAAGCTTTTCTGAACTCTTACTTCCATCTGTGGGACAGATTCCTCTCTTAT	540
QY	559	CTAGAACACCTTCAGGCATATCCTTATCTCTTTAGCAGACTCACTCTGCTGCTATTTAGC	618
Db	541	CTAGAACACCTTCAGGCATATCCTTATCTCTTTAGCAGACTCACTCTGCTGCTATTTAGC	600
QY	619	TCTCAGATTACATGTCCTTTCCCTACA	645
Db	601	TCTCAGATTACATGTCCTTTCCCTACA	627
RESULT 2			
BG676811			
LOCUS			
DEFINITION			
602623406F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4748664 5',			
mRNA sequence.			
ACCESSION			
BG676811			
VERSION			
BG676811.1 GI:13908208			
KEYWORDS			
EST.			
SOURCE			
Homo sapiens (human)			

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 778)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rsb@nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10600 row: k column: 01
High quality sequence stop: 777.
Location/Qualifiers
1. .778
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4748664"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"
/notes="Organ: Skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match		24.9%; Score 617.6; DB 12; Length 778;	
Best Local Similarity		96.2%; Pred. No. 6.9e-109;	
Matches	750; Conservative	0; Mismatches 19; Indels 11; Gaps 11;	
QY	1208	AAGGATGTTCTAATCTTTCTGCTCTGAGACGAATGCTATGGCTGCAGATGACTTCTTA 1267	
Db	3	AAGGATGTTCTAATCTTTCTGCTCTGCTGCTGAGACGAATGCTATGGCTGCAGATGACTTCTTA 62	
QY	1268	GAGGATTTGCTTTTGAGCAATAGGTAGATGGTTGGTGGTGT-GGAAAGCTTGGAGAG 1326	
Db	63	GAGGATTTGCTTTTGAGCATATAGGTAGATGGTCTGGGGGTGTAGAAAACCTTGGAGAG 122	
QY	1327	GTCAGGTAGTTGG-CTACTTTCTGTTGGATCTATTAAATACCTGGCAGCTCTCTGTCTT 1385	
Db	123	GTCAGGTAGTTGAGTACTTTCTGCTTGGATCTATTAAATATCTGGCAGACTCTCTGTCT 182	
QY	1386	TTTGTGGTGTGTCCTGTGATGATGCTGCTTTTAAACCACTCCCTCGATGCA-TTT 1444	
Db	183	TTTGTGGTGTGTCCTGTGATGATGCTGCTTTTAAACCACTCCCTCGATGCACTTC 242	
QY	1445	TTCCCTCCTTGCAATTCCTCTTTTCTGAGTTCTACTAGAGAAATCTGCACCTATGTTT 1504	
Db	243	TTCCCTCCTTGCAATTCCTCTTTTCTGAGTTCTACTAGAGAAATCTGCACCTATGTTT 302	
QY	1505	TTCCCTTTTGTCTGAGATGAAAGTTTAAATAATCCACCTCTGTCAATTCACCTCTC 1564	
Db	303	TTCCC-TTTTGTCTTGAGATGAAAGTTTAAATAATCCACCTCTGTCAATTCACCTCTC 361	
QY	1565	TGAACATCCCAAGCTGATCCCTGGCTCTTTCTCAGACTATGTTTCTTACTTGGGA- 1623	
Db	362	TGAACATCCCAAGCTGATCCCTGGCTCTTTCTCAGACTATGTTTCTTACTTGGGAG 421	
QY	1624	CCTAGAACTGGATTTGGCATTTGCTCCTGATCAGATGAGACCTTTGATTTTGGCCC 1683	
Db	422	CCTAGAACTGGATTTGGCATTTGCTCCTGATCAGATGAGACCTTTGATTTTGGCCC 481	
QY	1684	CTTCTTTAGGACCTTACACT-CCTGTCTTTCTTTGACTTGGCTTTTGTCTTCTTCTTC 1742	
Db	482	CTTCTTTAGGACCTTACACTTCTTCTTTGACTTGGCTTTTGTCTTCTTCTTCTTC 541	

```

QY 1743 ATCTTAGTCCCTCTTCATGCGATATGGTCATTGCTA-GGTAGAGGTATGCTCTTTATGT 1801
DB 542 ATCTTAGTCCCTCTTCATGCGATATGGTCATTGCTAGGTAGAGGTATGCTCTTTATGT 601
QY 1802 AATGGCCACCCGCAATTAGTATTACATAAA-CTTTCTTTTAAACAATCTGTGCAATGACAT 1860
DB 602 AATGGCCACCCGCAATTAGTATTACATAAAAGCTTTCTTTTAAACAATCTGTGCAATGACAT 661
QY 1861 GCTGCTCTGTTCCATTTAGAGATTGACAGAGGTTTCAGTTTGTATGATCTCAATCTTAT 1920
DB 662 GCTGCTCTGTTCCATTTAGAGATTGACAGAGG-TTCAGTTTGTATGATCTCAATCTTAT 718
QY 1921 TTTAGTCTTGGGAAATCAATTCAGAAATACATCTCTCCAAATCTCTCTTACTCAAA 1980
DB 719 TTTAGTCTTGGGAAATCAATTCAGAAATACATCTCTCCAAATCTCTCTTACTCAAA 778

RESULT 3
AA180763 607 bp mRNA linear EST 10-MAR-1998
LOCUS zp43f01.v1 Stratagene muscle 937209 Homo sapiens cDNA clone
DEFINITION IMAGE:612217.5, similar to SW.P44 PANTR P27473 NON-A NON-B
HEPATITIS-ASSOCIATED MICROTUBULAR AGGREGATES PROTEIN ; mRNA
sequence.
ACCESSION AA180763
VERSION AA180763
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 607)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1453 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 386.

FEATURES
source
1..607
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:4644056"
/db_xref="taxon:9606"
/clone="IMAGE:612217"
/tissue_type="muscle"
/dev_stage="adult"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene muscle 937209"
/note="Organ: skeletal muscle; Vector: pBluescript SK-;
Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally.
Primer: Oligo dt. Skeletal muscle from patient with
malignant hyperthermia. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG
3' -3' adaptor sequence: 5' CTCGAGTATTTTCTTTTCTTACTTATTC 1182

QY 1123 GGAAGTCCAGAAACTGGATTGCTCTTCTGACATCTCGGTGGTTAGCAATTATTC 1182

```

```

DB 1 GGAAGTCCAGAAACTGGATTGCTCTTCTGACATCTCGGTGGTTAGCAATTATTC 60
QY 1183 CTCTGAGTGGGAGCTGGAGCCCTGTAAAGAGATGTTCTAAATCTTCTGCTCGAGACGAT 1242
DB 61 CTCTGAGTGGGAGCTGGAGCCCTGTAAAGAGATGTTCTAAATCTTCTGCTCGAGACGAT 120
QY 1243 GCTATGGGCTGAGAGTCTCTTCTAGAGGATTTGCTTTGAGCAAAATAGGTAGATGGTT 1302
DB 121 GCTATGGGCTGAGAGTCTCTTCTAGAGGATTTGCTTTTGGCAAAATAGGTAGATGGTT 180
QY 1303 TGCTGTGTGGAAGCTTGGAGCGGTGAGGTAGTTGGCTACTTTCTGCTTGGATCTATTA 1362
DB 181 TGCTGTGTGGAAGCTTGGAGCGGTGAGGTAGTTGGCTACTTTCTGCTTGGATCTATTA 240
QY 1363 AATACCTGCGACCTCTCTGCTCTTTTGGGTGTTTGGCCCTGATAGTTCTGCTTTT 1422
DB 241 AATACCTGCGACCTCTCTGCTCTTTTGGGTGTTTGGCCCTGATAGTTCTGCTTTT 300
QY 1423 AACCCACTCCCTGGATGCAATTTTCCCTCTTCTGCAATTTCCCTCTTTTCTCGAGTTCATA 1482
DB 301 AACCCACTCCCTGGATGCAATTTTCCCTCTTCTGCAATTTCCCTCTTTTCTCGAGTTCATA 360
QY 1483 CTAGAGAAATCTGCATATGTTTTTCCCTTTTCTGCTTGGATGAAAGTTTAAATAATC 1542
DB 361 CTAGAGAAATCTGCATATGTTTTTCCCTTTTCTGCTTGGATGAAAGTTTAAATAATC 420
QY 1543 CACCTCTGTCAATTTCCACTCTCTGAACATCCCAAGCTGTATCCCTGGCCCTTTTCTCAG 1602
DB 421 CACCTCTGTCAATTTCCACTCTCTGAACATCCCAAGCTGTAT-CTTGGCCCTTTTCTCAG 479
QY 1603 ACTATGTTTTCTTACTTGGGACCTAGAACTGGATGGATGGCAATGCTCTCTGATCAGAT 1662
DB 480 ACTAAGTTTCTTACTTGGGACCTAGAACTGGATGGATGGCA-TGCTCCNGATCAGAT 538
QY 1663 GAGACCTTTGATTTATTTGGCCCTTCTCTAGGACCTTACACTCTCTTCTTCTTCTGACTTG 1722
DB 539 GAGACCTT---GAATATTGCCCTCTCTTAGACCTTA-ANTCCCGCCTTCTTCTGACTGC 594
QY 1723 CCTTT 1727
DB 595 CTTTT 599

RESULT 4
AA480136 513 bp mRNA linear EST 08-AUG-1997
LOCUS zv42f09.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
DEFINITION IMAGE:756329.5, mRNA sequence.
ACCESSION AA480136
VERSION AA480136
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 513)
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-Merck EST Project 1997
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 ET from Amersham
High quality sequence stop: 493.

REFERENCE
AUTHORS
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-Merck EST Project 1997
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 ET from Amersham
High quality sequence stop: 493.

```

```
FEATURES
source
  Location/Qualifiers
    1..513
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:756329"
      /sex="Female"
      /tissue_type="ovarian tumor"
      /lab_host="DH10B (ampicillin resistant)"
      /clone_lib="Soares ovary tumor NBHOT"
      /notes="Organ: ovary; Vector: p7T3D (Pharmacia) with a
      modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
      strand cDNA was primed with a Not I - oligo(dT) primer [5',
      TGTACCAATCTGAGTGGAGCGCGCTTTTATTTTATTTT 3'],
      double-stranded cDNA was size selected, ligated to Eco RI
      adapters (Pharmacia), digested with Not I and cloned into
      the Not I and Eco RI sites of a modified p7T3 vector
      (Pharmacia). Library constructed by Bento Soares and
      M.Fatima Bonaldo."
  ORIGIN
    Query Match      20.7%; Score 513; DB 9; Length 513;
    Best Local Similarity 100.0%; Pred. No. 1.1e-88;
    Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 1650 CTCCTGATCAGATGAGACCTTTGATTTATTTGCCCTTCCTTAGGACCTTACACTCTCTGTC 1709
  Db 1 CTCCTGATCAGATGAGACCTTTGATTTATTTGCCCTTCCTTAGGACCTTACACTCTCTGTC 60
  QY 1710 TTTCTTTGACTTGGCTTTTGTCTTCTTCATCTAGTCCCTTCATGCAATGCG 1769
  Db 61 TTTCTTTGACTTGGCTTTTGTCTTCTTCATCTAGTCCCTTCATGCAATGCG 120
  QY 1770 TCATTGCTAGGTAGAGGTATGCTCTTTATGTAATGCCACCGCATTTAGTATTACATAA 1829
  Db 121 TCATTGCTAGGTAGAGGTATGCTCTTTATGTAATGCCACCGCATTTAGTATTACATAA 180
  QY 1830 ACTTTCTTTTACAACTGTCATGATGCTGCTCTGTTCCATTTAGAGATTGACA 1889
  Db 181 ACTTTCTTTTAACAATCTGTCATGATGCTGCTGTTCCATTTAGAGATTGACA 240
  QY 1890 GAGGTTTCAGTTTAGTATACTCAAACTTTATTTTAGTGTGCGGAAATCAATTCAGAATA 1949
  Db 241 GAGGTTTCAGTTTAGTATACTCAAACTTTATTTTAGTGTGCGGAAATCAATTCAGAATA 300
  QY 1950 TCATCTCTCTCCAACTCTCTTACTCAAACTGTCGGAAACTCTCATGTTACTAACTT 2009
  Db 301 TCATCTCTCTCCAACTCTCTTACTCAAACTGTCGGAAACTCTCATGTTACTAACTT 360
  QY 2010 TGTGTCTTAACCTGCTGCTTGTGTTTCCCATCCCTTCTCTTCTCATGTCGTGTG 2069
  Db 361 TGTGTCTTAACCTGCTGCTTGTGTTTCCCATCCCTTCTCTTCTCATGTCGTGTG 420
  QY 2070 CTCCTTAATATAGCGTTGGTGGATTTTCAGTGGTCCAAATATTCCTCTCCCTCTGGTT 2129
  Db 421 CTCCTTAATATAGCGTTGGTGGATTTTCAGTGGTCCAAATATTCCTCTCCCTCTGGTT 480
  QY 2130 GCCTTTCTCAGATAATCCACTAAGAAATTTT 2162
  Db 481 GCCTTTCTCAGATAATCCACTAAGAAATTTT 513
  RESULT 5
  BE049439/c
  LOCUS
  DEFINITION
    xx86e11.x1 NCI_CGAP_Panl Homo sapiens cDNA clone IMAGE:2834924 3',
    mRNA sequence.
  ACCESSION
    BE049439
  VERSION
    BE049439.1
  KEYWORDS
    EST.
  SOURCE
    Homo sapiens (human)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 510)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-i@mail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL, send email to:
info@image.llnl.gov
Possible reversed clone: polyT not found
Seq primer: -40UP from Gibco
High quality sequence stop: 412.
Location/Qualifiers
  1..510
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:2834924"
    /tissue_type="adenocarcinoma"
    /lab_host="DH10B"
    /clone_lib="NCI_CGAP_Panl"
    /notes="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: SalI;
    Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
    Average insert size 1.72 kb. Life Technologies catalog #:
    11548-013"
  ORIGIN
    Query Match      20.5%; Score 510; DB 10; Length 510;
    Best Local Similarity 100.0%; Pred. No. 4.1e-88;
    Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 1970 TCTTACTCAAAATGCTGGGAACTCTCATGTTACTTAACCTTTGTGCTCTAACTCTGCCAT 2029
  Db 510 TCTTACTCAAAATGCTGGGAACTCTCATGTTACTTAACCTTTGTGCTCTAACTCTGCCAT 451
  QY 2030 CTGGTTTCCCATCCCTTCTCTCTCATGTTAGTGTGCTCTTAATATTAGCTGGT 2089
  Db 450 CTGGTTTCCCATCCCTTCTCTCTCATGTTAGTGTGCTCTTAATATTAGCTGGT 391
  QY 2090 TGAGATTTTCAGTGTGCTCAATATTCCTCTCTCTGCTGCTCTTCTCTGAGATAATCCA 2149
  Db 390 TGAGATTTTCAGTGTGCTCAATATTCCTCTCTCTGCTGCTCTTCTCTGAGATAATCCA 331
  QY 2150 CTAAGATAATTTGTGTTCTTTTCTCAGGGAATCTAAGGAGGAAATATCAACTGTGC 2209
  Db 330 CTAAGATAATTTGTGTTCTTTTCTCAGGGAATCTAAGGAGGAAATATCAACTGTGC 271
  QY 2210 ACAAGGAAAAATAGATATGTGAAGGTTCACTAAATTTCTCTCATCACAGAGATT 2269
  Db 270 ACAAGGAAAAATAGATATGTGAAGGTTCACTAAATTTCTCTCATCACAGAGATT 211
  QY 2270 AAAATT CAGAAAGGAGAAAAACACAGACCAAGAGAGATATCTAAGACCAAGGATGTGT 2329
  Db 210 AAAATT CAGAAAGGAGAAAAACACAGACCAAGAGAGATATCTAAGACCAAGGATGTGT 151
  QY 2330 TTTTATTATGCTAGGATGAGAAATGCATAGAACATTTAGTACTTTGTAATAACTAGA 2389
  Db 150 TTTTATTATGCTAGGATGAGAAATGCATAGAACATTTAGTACTTTGTAATAACTAGA 91
  QY 2390 AATAACATGATTTAGTCATTAATTTGTGAAAAATATAATAATTTTCTTGGATTATGTTTC 2449
  Db 90 AATAACATGATTTAGTCATTAATTTGTGAAAAATATAATAATTTTCTTGGATTATGTTTC 31
  QY 2450 TGTATCTGTGAAAAATATAATAATTTCTTATAA 2479
  Db 30 TGTATCTGTGAAAAATATAATAATTTCTTATAA 1
  RESULT 6
```

```

N90350      639 bp      mRNA      linear      EST 03-APR-1996
LOCUS       Yz88e07.r1 Soares multiple sclerosis 2NBHMP Homo sapiens cDNA
DEFINITION  Clone IMAGE:290148 5', mRNA sequence.
ACCESSION   N90350
VERSION     N90350.1 GI:1443677
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 639)
AUTHORS    Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterson,A., Williamson,A., Wohldmann,P. and
Wilson,R.
TITLE      The WashU-Merck EST Project
JOURNAL     Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: Enprimer
            High quality sequence stop: 408.
FEATURES    Location/Qualifiers
             ..639
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="GDB:3905988"
             /db_xref="taxon:9606"
             /clone="IMAGE:290148"
             /sex="male"
             /tissue_type="multiple sclerosis lesions"
             /dev_stage="Age 46"
             /lab_host="DH10B (ampicillin resistant)"
             /clone_lib="Soares multiple sclerosis 2NBHMP"
             /notes="Vector: pT73D (Pharmacia) with a modified
polylinker V.TPE; phagemid; Site 1: Not I; Site 2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5'
TGTTACCAATCTGAAGTGGAGCGGCCGCAATTTTTTTTTTTTTTTT 3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA from 4 multiple sclerosis
lesions from one patient was kindly provided by Dr. Kevin
G. Becker (NINDS/NIH)."
ORIGIN
Query Match      20.3%; Score 504; DB 14; Length 639;
Best Local Similarity 92.3%; Pred.No. 5.6e-87;
Matches 560; Conservative 0; Mismatches 35; Indels 8; Gaps 3;

QY      1821  ATTACATAAACTTTCTTTTAAACAATCTGTGCATGACATGCTCTCTTCATTAGA 1880
DB      1      ATTACATAAACTTTCTTTTAAACAATCTGTGCATGACATGCTCTCTTCATTAGA 60
QY      1881  GATTGACAGAGTTTCAGTTTGTAGTACTCAATCTTATTTAGTGTGGGAATCAA 1940
DB      61      GATTGACAGAGTTTCAGTTTGTAGTACTCAATCTTATTTAGTGTGGGAATCAA 120
QY      1941  TTGAGAAATACATCTCTCCAAATCTCTCTTACTCAAAATGCTGGAAATCTCATCT 2000
DB      121      TTGAGAAATACATCTCTCCAAATCTCTCTTACTCAAAATGCTGGAAATCTCATCT 180
QY      2001  TACTAACTTTGGTGTCTAACTCTGCCATCTTGGTTCCCATCCCTTCTCTTCCTCATG 2060

```

```

181  TACTAACTTTGGTGTCTAACTCTGCCATCTTGGTTCCCATCCCTTCTCTTCCTCATG 240
2061 GTACGTGTGCTCTTAATATTAGCGTTGGTGGAGATTTTCAGTGGTCAATATTCTCTTC 2120
241  GTACGTGTGCTCTTAATATTAGCGTTGGTGGAGATTTTCAGTGGTCAATATTCTCTTC 300
2121 CCTCTGGTGGCTTTCTCTGAGATAATCCACTAGATAATTTTGTGTCTTTTCTCAGGG 2180
301  CCTCTGGTGGCTTTCTCTGAGATAATCCACTAGATAATTTTGTGTCTTTTCTCTC 357
2181 AATCTAGGAGGAGAAATTTATCACTGTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 2240
358 GAATTAAGGAGGAGAAATTTATCACTGTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 417
2241 ACCTAAATTTCTCTACATCAAGAGATTAATTCAGAAAGGAGGAGGAGGAGGAGGAGGAGG 2300
418 ACCTAAATTTCTCTACATCAAGAGATTAATTCAGAAAGGAGGAGGAGGAGGAGGAGGAGG 477
2301 GAGAAGTATCTAAGACCAAGGGATGTGTTTATTATTTAGTCTAGGATGAAGAAATGCATA 2360
478 GAGAAGTANTAAAGACCAAGGGATGTG-TTTATTATGCTAGGATGAAGAAATGCATA 536
2361 GAACATTTAGTACTGTGTAATAACTAGAAATACATGATTAGTCTATATTGTGAAAAA 2420
537 GA-----CATTTAGTCTTGTAACTACTAGAAATACCATGATTAGGCCATAATGTGAAAAA 592
2421 TAA 2423
593 AAA 595

RESULT 7
AA031928/c 504 bp mRNA linear EST 09-MAY-1997
LOCUS      IMAGE.470583 3', mRNA sequence.
DEFINITION AA031928
ACCESSION  AA031928
VERSION    AA031928.1 GI:1501891
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE   1 (bases 1 to 504)
AUTHORS    Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,
Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.,
and Marra,M.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 504)
Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,
Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.,
and Marra,M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
PUBMED     8889549
COMMENT    Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 670 Std Error: 0.00
            Seq primer: -40M13 fwd. from Amersham
            High quality sequence stop: 446.
            Location/Qualifiers
             ..504
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="GDB:3756585"
             /db_xref="taxon:9606"
FEATURES    source

```